Noname manuscript No.
(will be inserted by the editor)

On the discrepancy principle for some Newton type methods for solving nonlinear inverse problems

Qinian Jin · **Ulrich Tautenhahn**

the date of receipt and acceptance should be inserted later

Abstract We consider the computation of stable approximations to the exact solution $x^\dagger$ of nonlinear ill-posed inverse problems $F(x) = y$ with nonlinear operators $F : X \to Y$ between two Hilbert spaces X and Y by the Newton type methods

$$x_{k+1}^\delta = x_0 - g_{\alpha_k}\left(F'(x_k^\delta)^*F'(x_k^\delta)\right)F'(x_k^\delta)^*\left(F(x_k^\delta) - y^\delta - F'(x_k^\delta)(x_k^\delta - x_0)\right)$$

in the case that only available data is a noise y^δ of y satisfying $\|y^\delta - y\| \le \delta$ with a given small noise level $\delta > 0$. We terminate the iteration by the discrepancy principle in which the stopping index k_δ is determined as the first integer such that

$$\|F(x_{k_\delta}^\delta) - y^\delta\| \le \tau\delta < \|F(x_k^\delta) - y^\delta\|, \qquad 0 \le k < k_\delta$$

with a given number $\tau > 1$. Under certain conditions on $\{\alpha_k\}$, $\{g_\alpha\}$ and F, we prove that $x_{k_\delta}^\delta$ converges to $x^\dagger$ as $\delta \to 0$ and establish various order optimal convergence rate results. It is remarkable that we even can show the order optimality under merely the Lipschitz condition on the Fréchet derivative F' of F if $x_0 - x^\dagger$ is smooth enough.

Keywords Nonlinear inverse problems · Newton type methods · the discrepancy principle · order optimal convergence rates

Mathematics Subject Classification (2000) 65J15 · 65J20 · 47H17

1 Introduction

In this paper we will consider the nonlinear inverse problems which can be formulated as the operator equations

$$F(x) = y, \tag{1.1}$$

where $F : D(F) \subset X \to Y$ is a nonlinear operator between the Hilbert spaces X and Y with domain $D(F)$. We will assume that problem (1.1) is ill-posed in the sense that its solution does not depend continuously on the right hand side y, which is the characteristic property for most of

Qinian Jin
Department of Mathematics, The University of Texas at Austin, Austin, Texas 78712, USA
E-mail: qjin@math.utexas.edu

Ulrich Tautenhahn
Department of Mathematics, University of Applied Sciences Zittau/Görlitz, PO Box 1454, 02754 Zittau, Germany
E-mail: u.tautenhahn@hs-zigr.de

the inverse problems. Such problems arise naturally from the parameter identification in partial differential equations.

Throughout this paper $\|\cdot\|$ and $(\cdot,\cdot)$ denote respectively the norms and inner products for both the spaces X and Y since there is no confusion. The nonlinear operator F is always assumed to be Fréchet differentiable, the Fréchet derivative of F at $x \in D(F)$ is denoted as $F'(x)$ and $F'(x)^*$ is used to denote the adjoint of $F'(x)$. We assume that y is attainable, i.e. problem (1.1) has a solution $x^\dagger \in D(F)$ such that

$$F(x^\dagger) = y.$$

Since the right hand side is usually obtained by measurement, thus, instead of y itself, the available data is an approximation y^δ satisfying

$$\|y^\delta - y\| \le \delta \tag{1.2}$$

with a given small noise level $\delta > 0$. Due to the ill-posedness, the computation of a stable solution of (1.1) from y^δ becomes an important issue, and the regularization techniques have to be taken into account.

Many regularization methods have been considered to solve (1.1) in the last two decades. Tikhonov regularization is one of the well-known methods that has been studied extensively (see [17, 11, 19] and the references therein). Due to the straightforward implementation, iterative methods are also attractive for solving nonlinear inverse problems. In this paper we will consider some Newton type methods in which the iterated solutions $\{x_k^\delta\}$ are defined successively by

$$x_{k+1}^\delta = x_0 - g_{\alpha_k}\left(F'(x_k^\delta)^*F'(x_k^\delta)\right)F'(x_k^\delta)^*\left(F(x_k^\delta) - y^\delta - F'(x_k^\delta)(x_k^\delta - x_0)\right), \tag{1.3}$$

where $x_0^\delta := x_0$ is an initial guess of $x^\dagger$, $\{\alpha_k\}$ is a given sequence of numbers such that

$$\alpha_k > 0, \qquad 1 \le \frac{\alpha_k}{\alpha_{k+1}} \le r \qquad \text{and} \qquad \lim_{k\to\infty} \alpha_k = 0 \tag{1.4}$$

for some constant $r > 1$, and $g_\alpha : [0,\infty) \to (-\infty,\infty)$ is a family of piecewise continuous functions satisfying suitable structure conditions. The method (1.3) can be derived as follows. Suppose x_k^δ is a current iterate, then we may approximate $F(x)$ by its linearization around x_k^δ, i.e. $F(x) \approx F(x_k^\delta) + F'(x_k^\delta)(x - x_k^\delta)$. Thus, instead of (1.1), we have the approximate equation

$$F'(x_k^\delta)(x - x_k^\delta) = y^\delta - F(x_k^\delta). \tag{1.5}$$

If $F'(x_k^\delta)$ has bounded inverse, the usual Newton method defines the next iterate by solving (1.5) for x. For nonlinear ill-posed inverse problems, however, $F'(x_k^\delta)$ in general is not invertible. Therefore, we must use linear regularization methods to solve (1.5). There are several ways to do this step. One way is to rewrite (1.5) as

$$F'(x_k^\delta)h = y^\delta - F(x_k^\delta) + F'(x_k^\delta)(x_k^\delta - x_0), \tag{1.6}$$

where $h = x - x_0$. Applying the linear regularization method defined by $\{g_\alpha\}$ we may produce the regularized solution h_k^δ by

$$h_k^\delta = g_{\alpha_k}\left(F'(x_k^\delta)^*F'(x_k^\delta)\right)F'(x_k^\delta)^*\left(y^\delta - F(x_k^\delta) + F'(x_k^\delta)(x_k^\delta - x_0)\right).$$

The next iterate is then defined to be $x_{k+1}^\delta := x_0 + h_k^\delta$ which is exactly the form (1.3).

In order to use x_k^δ to approximate $x^\dagger$, we must choose the stopping index of iteration properly. Some Newton type methods that can be casted into the form (1.3) have been analyzed in [3, 12, 14] under a priori stopping rules, which, however, depend on the knowledge of the smoothness of $x_0 - x^\dagger$ that is difficult to check in practice. Thus a wrong guess of the smoothness will lead to a bad choice of the stopping index, and consequently to a bad approximation to $x^\dagger$. Therefore,

a posteriori rules, which use only quantities that arise during calculations, should be considered to choose the stopping index of iteration. One can consult [3,8,4,9,2,14] for several such rules.

One widely used a posteriori stopping rule in the literature of regularization theory for ill-posed problems is the discrepancy principle which, in the context of the Newton method (1.3), defines the stopping index k_δ to be the first integer such that

$$\|F(x_{k_\delta}^\delta)-y^\delta\| \le \tau\delta < \|F(x_k^\delta)-y^\delta\|, \quad 0\le k<k_\delta, \tag{1.7}$$

where $\tau>1$ is a given number. The method (1.3) with $g_\alpha(\lambda)=(\alpha+\lambda)^{-1}$ together with (1.7) has been considered in [3,8]. Note that when $g_\alpha(\lambda)=(\alpha+\lambda)^{-1}$, the method (1.3) is equivalent to the iteratively regularized Gauss-Newton method [1]

$$x_{k+1}^\delta = x_k^\delta - \Big(\alpha_k I + F'(x_k^\delta)^* F'(x_k^\delta)\Big)^{-1}\Big(F'(x_k^\delta)^*(F(x_k^\delta)-y^\delta)+\alpha_k(x_k^\delta-x_0)\Big). \tag{1.8}$$

When F satisfies the condition like

$$\begin{aligned} F'(x) &= R(x,z)F'(z)+Q(x,z),\\ \|I-R(x,z)\| &\le C_R\|x-z\|, \qquad\qquad x,z\in B_\rho(x^\dagger),\\ \|Q(x,z)\| &\le C_Q\|F'(z)(x-z)\|, \end{aligned} \tag{1.9}$$

where C_R and C_Q are two positive constants, for the method defined by (1.8) and (1.7) with τ being sufficiently large, it has been shown in [3,8] that if $x_0-x^\dagger$ satisfies the Hölder source condition

$$x_0-x^\dagger = (F'(x^\dagger)^*F'(x^\dagger))^\nu\omega \tag{1.10}$$

for some $\omega\in X$ and $0\le\nu\le 1/2$, then

$$\|x_{k_\delta}^\delta - x^\dagger\| \le o(\delta^{2\nu/(1+2\nu)});$$

while if $x_0-x^\dagger$ satisfies the logarithmic source condition

$$x_0-x^\dagger = \big(-\log(F'(x^\dagger)^*F'(x^\dagger))\big)^{-\mu}\omega \tag{1.11}$$

for some $\omega\in X$ and $\mu>0$, then

$$\|x_{k_\delta}^\delta - x^\dagger\| \le O((-\ln\delta)^{-\mu}).$$

Unfortunately, except the above results, there is no more result available in the literature on the general method defined by (1.3) and (1.7).

During the attempt of proving regularization property of the general method defined by (1.3) and (1.7), Kaltenbacher realized that the arguments in [3,8] depend heavily on the special properties of the function $g_\alpha(\lambda)=(\alpha+\lambda)^{-1}$, and thus the technique therein is not applicable. Instead of the discrepancy principle (1.7), she proposed in [13] a new a posteriori stopping rule to terminate the iteration as long as

$$\max\Big\{\|F(x_{m_\delta-1}^\delta)-y^\delta\|, \|F(x_{m_\delta-1}^\delta)+F'(x_{m_\delta-1}^\delta)(x_{m_\delta}^\delta-x_{m_\delta-1}^\delta)-y^\delta\|\Big\}\le\tau\delta \tag{1.12}$$

is satisfied for the first time, where $\tau>1$ is a given number. Under the condition like (1.9), it has been shown that if $x_0-x^\dagger$ satisfies the Hölder source condition (1.10) for some $\omega\in X$ and $0\le\nu\le 1/2$, then there hold the order optimal convergence rates

$$\|x_{m_\delta}^\delta - x^\dagger\| \le C_\nu\|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}$$

if $\{g_\alpha\}$ satisfies some suitable structure conditions, τ is sufficiently large and $\|\omega\|$ is sufficiently small. Note that any result on (1.12) does not imply that the corresponding result holds for (1.7). Note also that $k_\delta\le m_\delta-1$ which means that (1.12) requires more iterations to be performed.

Moreover, the discrepancy principle (1.7) is simpler than the stopping rule (1.12). Considering the fact that it is widely used in practice, it is important to give further investigations on (1.7).

In this paper, we will resume the study of the method defined by (1.3) and (1.7) with completely different arguments. With the help of the ideas developed in [9,19,10], we will show that, under certain conditions on $\{g_\alpha\}$, $\{\alpha_k\}$ and F, the method given by (1.3) and (1.7) indeed defines a regularization method for solving (1.1) and is order optimal for each $0 < \nu \le \bar{\nu} - 1/2$, where $\bar{\nu} \ge 1$ denotes the qualification of the linear regularization method defined by $\{g_\alpha\}$. In particular, when $x_0 - x^\dagger$ satisfies (1.10) for $1/2 \le \nu \le \bar{\nu} - 1/2$, we will show that the order optimality of (1.3) and (1.7) even holds under merely the Lipschitz condition on F'. This is the main contribution of the present paper. We point out that our results are valid for any $\tau > 1$. This less restrictive requirement on τ is important in numerical computations since the absolute error could increase with respect to τ.

This paper is organized as follows. In Section 2 we will state various conditions on $\{g_\alpha\}$, $\{\alpha_k\}$ and F, and then present several convergence results on the methods defined by (1.3) and (1.7). We then complete the proofs of these main results in Sections 3, 4, and 5. In Section 6, in order to indicate the applicability of our main results, we verify those conditions in Section 2 for several examples of $\{g_\alpha\}$ arising from Tikhonov regularization, the iterated Tikhonov regularization, the Landweber iteration, the Lardy's method, and the asymptotic regularization.

2 Assumptions and main results

In this section we will state the main results for the method defined by (1.3) and the discrepancy principle (1.7). Since the definition of $\{x_k^\delta\}$ involves F, g_α and $\{\alpha_k\}$, we need to impose various conditions on them.

We start with the assumptions on g_α which is always assumed to be continuous on $[0,1/2]$ for each $\alpha > 0$. We will set

$$r_\alpha(\lambda) := 1 - \lambda g_\alpha(\lambda),$$

which is called the residual function associated with g_α.

Assumption 1 [1] *(a) There are positive constants c_0 and c_1 such that*

$$0 < r_\alpha(\lambda) \le 1, \quad r_\alpha(\lambda)\lambda \le c_0\alpha \quad \textit{and} \quad 0 \le g_\alpha(\lambda) \le c_1\alpha^{-1}$$

for all $\alpha > 0$ and $\lambda \in [0,1/2]$;

(b) $r_\alpha(\lambda) \le r_\beta(\lambda)$ for any $0 < \alpha \le \beta$ and $\lambda \in [0,1/2]$;

(c) There exists a constant $c_2 > 0$ such that

$$r_\beta(\lambda) - r_\alpha(\lambda) \le c_2 \sqrt{\frac{\lambda}{\alpha}} r_\beta(\lambda)$$

for any $0 < \alpha \le \beta$ and $\lambda \in [0,1/2]$.

The conditions (a) and (b) in Assumption 1 are standard in the analysis of linear regularization methods. Assumption 1(a) clearly implies

$$0 \le r_\alpha(\lambda)\lambda^{1/2} \le c_3\alpha^{1/2} \quad \text{and} \quad 0 \le g_\alpha(\lambda)\lambda^{1/2} \le c_4\alpha^{-1/2} \tag{2.1}$$

with $c_3 \le c_0^{1/2}$ and $c_4 \le c_1^{1/2}$. We emphasize that direct estimates on $r_\alpha(\lambda)\lambda^{1/2}$ and $g_\alpha(\lambda)\lambda^{1/2}$ could give smaller c_3 and c_4. From Assumption 1(a) it also follows for each $0 \le \nu \le 1$ that $r_\alpha(\lambda)\lambda^\nu \le c_0^\nu \alpha^\nu$ for all $\alpha > 0$ and $\lambda \in [0,1/2]$. Thus the linear regularization method defined by $\{g_\alpha\}$ has qualification $\bar{\nu} \ge 1$, where, according to [20], the qualification is defined to be the

[1] Recently we realized that (c) can be derived from (a) and (b).

largest number $\bar{\nu}$ with the property that for each $0 \le \nu \le \bar{\nu}$ there is a positive constant d_ν such that

$$r_\alpha(\lambda)\lambda^\nu \le d_\nu \alpha^\nu \quad \text{for all } \alpha > 0 \text{ and } \lambda \in [0,1/2]. \tag{2.2}$$

Moreover, Assumption 1(a) implies for every $\mu > 0$ that

$$r_\alpha(\lambda)(-\ln\lambda)^{-\mu} \le \min\left\{(-\ln\lambda)^{-\mu}, c_0\alpha\lambda^{-1}(-\ln\lambda)^{-\mu}\right\}$$

for all $0 < \alpha \le \alpha_0$ and $\lambda \in [0,1/2]$. It is clear that $(-\ln\lambda)^{-\mu} \le (-\ln(\alpha/(2\alpha_0)))^{-\mu}$ for $0 \le \lambda \le \alpha/(2\alpha_0)$. By using the fact that the function $\lambda \to c_0\alpha\lambda^{-1}(-\ln\lambda)^{-\mu}$ is decreasing on the interval $(0, e^{-\mu}]$ and is increasing on the interval $[e^{-\mu}, 1)$, it is easy to show that there is a positive constant a_μ such that $c_0\alpha\lambda^{-1}(-\ln\lambda)^{-\mu} \le a_\mu(-\ln(\alpha/(2\alpha_0)))^{-\mu}$ for $\alpha/(2\alpha_0) \le \lambda \le 1/2$. Therefore for every $\mu > 0$ there is a positive constant b_μ such that

$$r_\alpha(\lambda)(-\ln\lambda)^{-\mu} \le b_\mu(-\ln(\alpha/(2\alpha_0)))^{-\mu} \tag{2.3}$$

for all $0 < \alpha \le \alpha_0$ and $\lambda \in [0,1/2]$. This inequality will be used to derive the convergence rate when $x_0 - x^\dagger$ satisfies the logarithmic source condition (1.11)

The condition (c) in Assumption 1 seems to appear here for the first time. It is interesting to note that one can verify it for many well-known linear regularization methods. Moreover, the conditions (b) and (c) have the following important consequence.

Lemma 1 *Under the conditions (b) and (c) in Assumption 1, there holds*

$$\|[r_\beta(A^*A) - r_\alpha(A^*A)]x\| \le \|\bar{x} - r_\beta(A^*A)x\| + \frac{c_2}{\sqrt{\alpha}}\|A\bar{x}\| \tag{2.4}$$

for all $x, \bar{x} \in X$, any $0 < \alpha \le \beta$ and any bounded linear operator $A : X \to Y$ satisfying $\|A\| \le 1/\sqrt{2}$.

Proof For any $0 < \alpha \le \beta$ we set

$$p_{\beta,\alpha}(\lambda) := \frac{r_\beta(\lambda) - r_\alpha(\lambda)}{r_\beta(\lambda)}, \qquad \lambda \in [0,1/2].$$

It follows from the conditions (a) and (b) in Assumption 1 that

$$0 \le p_{\beta,\alpha}(\lambda) \le \min\left\{1, c_2\sqrt{\frac{\lambda}{\alpha}}\right\}. \tag{2.5}$$

Therefore, for any $x, \bar{x} \in X$,

$$\begin{aligned} \|[r_\beta(A^*A) - r_\alpha(A^*A)]x\| &= \|p_{\beta,\alpha}(A^*A)r_\beta(A^*A)x\| \\ &\le \|p_{\beta,\alpha}(A^*A)[r_\beta(A^*A)x - \bar{x}]\| + \|p_{\beta,\alpha}(A^*A)\bar{x}\| \\ &\le \|r_\beta(A^*A)x - \bar{x}\| + \|p_{\beta,\alpha}(A^*A)\bar{x}\|. \end{aligned} \tag{2.6}$$

Let $\{E_\lambda\}$ be the spectral family generated by A^*A. Then it follows from (2.5) that

$$\begin{aligned} \|p_{\beta,\alpha}(A^*A)\bar{x}\|^2 &= \int_0^{1/2} \left[p_{\beta,\alpha}(\lambda)\right]^2 d\|E_\lambda \bar{x}\|^2 \\ &\le c_2^2 \int_0^{1/2} \frac{\lambda}{\alpha} d\|E_\lambda \bar{x}\|^2 = \frac{c_2^2}{\alpha}\|(A^*A)^{1/2}\bar{x}\|^2 \\ &= \frac{c_2^2}{\alpha}\|A\bar{x}\|^2. \end{aligned}$$

Combining this with (2.6) gives the desired assertion. □

For the sequence of positive numbers $\{\alpha_k\}$, we will always assume that it satisfies (1.4). Moreover, we need also the following condition on $\{\alpha_k\}$ interplaying with r_α.

Assumption 2 *There is a constant $c_5 > 1$ such that*

$$r_{\alpha_k}(\lambda) \le c_5 r_{\alpha_{k+1}}(\lambda)$$

for all k and $\lambda \in [0, 1/2]$.

We remark that for some $\{g_\alpha\}$ Assumption 2 is an immediate consequence of (1.4). However, this is not always the case; in some situations, Assumption 2 indeed imposes further conditions on $\{\alpha_k\}$. As a rough interpretation, Assumption 2 requires for any two successive iterated solutions the errors do not decrease dramatically. This may be good for the stable numerical implementations of ill-posed problems although it may require more iterations to be performed. Note that Assumption 2 implies

$$\|r_{\alpha_k}(A^*A)x\| \le c_5 \|r_{\alpha_{k+1}}(A^*A)x\| \tag{2.7}$$

for any $x \in X$ and any bounded linear operator $A : X \to Y$ satisfying $\|A\| \le 1/\sqrt{2}$.

Throughout this paper, we will always assume that the nonlinear operator $F : D(F) \subset X \to Y$ is Fréchet differentiable such that

$$B_\rho(x^\dagger) \subset D(F) \qquad \text{for some } \rho > 0 \tag{2.8}$$

and

$$\|F'(x)\| \le \min\left\{c_3 \alpha_0^{1/2}, \beta_0^{1/2}\right\}, \qquad x \in B_\rho(x^\dagger), \tag{2.9}$$

where $0 < \beta_0 \le 1/2$ is a number such that $r_{\alpha_0}(\lambda) \ge 3/4$ for all $\lambda \in [0, \beta_0]$. Since $r_{\alpha_0}(0) = 1$, such β_0 always exists. The scaling condition (2.9) can always be fulfilled by rescaling the norm in Y.

The convergence analysis on the method defined by (1.3) and (1.7) will be divided into two cases:

(i) $x_0 - x^\dagger$ satisfies (1.10) for some $\nu \ge 1/2$;
(ii) $x_0 - x^\dagger$ satisfies (1.10) with $0 \le \nu < 1/2$ or (1.11) with $\mu > 0$.

Thus different structure conditions on F will be assumed in order to carry out the arguments. It is remarkable to see that for case (i) the following Lipschitz condition on F' is enough for our purpose.

Assumption 3 *There exists a constant L such that*

$$\|F'(x) - F'(z)\| \le L\|x - z\| \tag{2.10}$$

for all $x, z \in B_\rho(x^\dagger)$.

As the immediate consequence of Assumption 3, we have

$$\|F(x) - F(z) - F'(z)(x - z)\| \le \frac{1}{2}L\|x - z\|^2$$

for all $x, z \in B_\rho(x^\dagger)$. We will use this consequence frequently in this paper.

During the convergence analysis of (1.3), we will meet some terms involving operators such as $r_{\alpha_k}(F'(x_k^\delta)^* F'(x_k^\delta))$. In order to make use of the source conditions (1.10) for $x_0 - x^\dagger$, we need to switch these operators with $r_{\alpha_k}(F'(x^\dagger)^* F'(x^\dagger))$. Thus we need the following commutator estimates involving r_α and g_α.

Assumption 4 *There is a constant $c_6 > 0$ such that*

$$\|r_\alpha(A^*A) - r_\alpha(B^*B)\| \le c_6 \alpha^{-1/2} \|A - B\|, \tag{2.11}$$

$$\|[r_\alpha(A^*A) - r_\alpha(B^*B)]B^*\| \le c_6 \|A - B\|, \tag{2.12}$$

$$\|A[r_\alpha(A^*A) - r_\alpha(B^*B)]B^*\| \le c_6 \alpha^{1/2} \|A - B\|, \tag{2.13}$$

and

$$\|[g_\alpha(A^*A) - g_\alpha(B^*B)]B^*\| \le c_6 \alpha^{-1} \|A - B\| \tag{2.14}$$

for any $\alpha > 0$ and any bounded linear operators $A, B : X \to Y$ satisfying $\|A\|, \|B\| \le 1/\sqrt{2}$.

This assumption looks restrictive. However, it is interesting to note that for several important examples we indeed can verify it easily, see Section 6 for details. Moreover, in our applications, we only need Assumption 4 with $A = F'(x)$ and $B = F'(z)$ for $x, z \in B_\rho(x^\dagger)$, which is trivially satisfied when F is linear.

Now we are ready to state the first main result of this paper.

Theorem 1 *Let $\{g_\alpha\}$ and $\{\alpha_k\}$ satisfy Assumption 1, (1.4), Assumption 2, and Assumption 4, let $\bar{\nu} \ge 1$ be the qualification of the linear regularization method defined by $\{g_\alpha\}$, and let F satisfy (2.8), (2.9) and Assumption 3 with $\rho > 4\|x_0 - x^\dagger\|$. Let $\{x_k^\delta\}$ be defined by (1.3) and let k_δ be the first integer satisfying (1.7) with $\tau > 1$. Let $x_0 - x^\dagger$ satisfy (1.10) for some $\omega \in X$ and $1/2 \le \nu \le \bar{\nu} - 1/2$. Then*

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\nu \|\omega\|^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)}$$

if $L\|u\| \le \eta_0$, where $u \in \mathcal{N}(F'(x^\dagger)^*)^\perp \subset Y$ is the unique element such that $x_0 - x^\dagger = F'(x^\dagger)^* u$, $\eta_0 > 0$ is a constant depending only on r, τ and c_i, and C_ν is a positive constant depending only on r, τ, ν and c_i, $i = 0, \cdots, 6$.

Theorem 1 tells us that, under merely the Lipschitz condition on F', the method (1.3) together with (1.7) indeed defines an order optimal regularization method for each $1/2 \le \nu \le \bar{\nu} - 1/2$; in case the regularization method defined by $\{g_\alpha\}$ has infinite qualification the discrepancy principle (1.7) provides order optimal convergence rates for the full range $\nu \in [1/2, \infty)$. This is one of the main contribution of the present paper.

We remark that under merely the Lipschitz condition on F' we are not able to prove the similar result as in Theorem 1 if $x_0 - x^\dagger$ satisfies weaker source conditions, say (1.10) for some $\nu < 1/2$. Indeed this is still an open problem in the convergence analysis of regularization methods for nonlinear ill-posed problems. In order to pursue the convergence analysis under weaker source conditions, we need stronger conditions on F than Assumption 3. The condition (1.9) has been used in [3,8] to establish the regularization property of the method defined by (1.8) and (1.7), where the special properties of $g_\alpha(\lambda) = (\lambda + \alpha)^{-1}$ play the crucial roles. In order to study the general method (1.3) under weaker source conditions, we need the following two conditions on F.

Assumption 5 *There exists a positive constant K_0 such that*

$$F'(x) = F'(z) R(x, z),$$
$$\|I - R(x, z)\| \le K_0 \|x - z\|$$

for any $x, z \in B_\rho(x^\dagger)$.

Assumption 6 *There exist positive constants K_1 and K_2 such that*

$$\|[F'(x) - F'(z)]w\| \le K_1 \|x - z\| \|F'(z) w\| + K_2 \|F'(z)(x - z)\| \|w\|$$

for any $x, z \in B_\rho(x^\dagger)$ and $w \in X$.

Assumption 5 has been used widely in the literature of nonlinear ill-posed problems (see [17,11,9,19]); it can be verified for many important inverse problems. Another frequently used assumption on F is (1.9) which is indeed quite restrictive. It is clear that Assumption 6 is a direct consequence of (1.9). In order to illustrate that Assumption 6 could be weaker than (1.9), we consider the identification of the parameter c in the boundary value problem

$$\begin{cases} -\Delta u + cu = f & \text{in } \Omega \\ u = g & \text{on } \partial\Omega \end{cases} \tag{2.15}$$

from the measurement of the state u, where $\Omega \subset \mathbb{R}^n, n \leq 3$, is a bounded domain with smooth boundary $\partial\Omega$, $f \in L^2(\Omega)$ and $g \in H^{3/2}(\partial\Omega)$. We assume $c^\dagger \in L^2(\Omega)$ is the sought solution. This problem reduces to solving an equation of the form (1.1) if we define the nonlinear operator F to be the parameter-to-solution mapping $F : L^2(\Omega) \to L^2(\Omega), F(c) := u(c)$ with $u(c) \in H^2(\Omega) \subset L^2(\Omega)$ being the unique solution of (2.15). Such F is well-defined on

$$D(F) := \left\{ c \in L^2(\Omega) : \|c - \hat{c}\|_{L^2} \leq \gamma \text{ for some } \hat{c} \geq 0 \text{ a.e.} \right\}$$

for some positive constant $\gamma > 0$. It is well-known that F has Fréchet derivative

$$F'(c)h = -A(c)^{-1}(hF(c)), \qquad h \in L^2(\Omega), \tag{2.16}$$

where $A(c) : H^2 \cap H_0^1 \to L^2$ is defined by $A(c)u := -\Delta u + cu$ which is an isomorphism uniformly in a ball $B_\rho(c^\dagger) \subset D(F)$ around $c^\dagger$. Let V be the dual space of $H^2 \cap H_0^1$ with respect to the bilinear form $\langle \varphi, \psi \rangle = \int_\Omega \varphi(x)\psi(x)dx$. Then $A(c)$ extends to an isomorphism from $L^2(\Omega)$ to V. Since (2.16) implies for any $c, d \in B_\rho(c^\dagger)$ and $h \in L^2(\Omega)$

$$\left(F'(c) - F'(d)\right) h = -A(c)^{-1}\left((c-d)F'(d)h\right) - A(c)^{-1}\left(h(F(c) - F(d))\right),$$

and since $L^1(\Omega)$ embeds into V due to the restriction $n \leq 3$, we have

$$\begin{aligned} \|(F'(c) - F'(d))h\|_{L^2} &\leq \|A(c)^{-1}\left((c-d)F'(d)h\right)\|_{L^2} + \|A(c)^{-1}\left(h(F(c) - F(d))\right)\|_{L^2} \\ &\leq C\|(c-d)F'(d)h\|_V + C\|h(F(c) - F(d))\|_V \\ &\leq C\|(c-d)F'(d)h\|_{L^1} + C\|h(F(c) - F(d))\|_{L^1} \\ &\leq C\|c-d\|_{L^2}\|F'(d)h\|_{L^2} + C\|F(c) - F(d)\|_{L^2}\|h\|_{L^2}. \end{aligned} \tag{2.17}$$

On the other hand, note that $F(c) - F(d) = -A(d)^{-1}((c-d)F(c))$, by using (2.16) we obtain

$$F(c) - F(d) - F'(d)(c-d) = -A(d)^{-1}((c-d)(F(c) - F(d))).$$

Thus, by a similar argument as above,

$$\|F(c) - F(d) - F'(d)(c-d)\|_{L^2} \leq C\|c-d\|_{L^2}\|F(c) - F(d)\|_{L^2}.$$

Therefore, if $\rho > 0$ is small enough, we have $\|F(c) - F(d)\|_{L^2} \leq C\|F'(d)(c-d)\|_{L^2}$, which together with (2.17) verifies Assumption 6. The validity of (1.9), however, requires $u(c) \geq \kappa > 0$ for all $c \in B_\rho(c^\dagger)$, see [7].

In our next main result, Assumption 5 and Assumption 6 will be used to derive estimates related to $x_k^\delta - x^\dagger$ and $F'(x^\dagger)(x_k^\delta - x^\dagger)$ respectively. Although Assumption 6 does not explore the full strength of (1.9), the plus of Assumption 5 could make our conditions stronger than (1.9) in some situations. One advantage of the use of Assumption 5 and Assumption 6, however, is that we can carry out the analysis on the discrepancy principle (1.7) for any $\tau > 1$, in contrast to those results in [3,8] where τ is required to be sufficiently large. It is not yet clear if only one of the above two assumptions is enough for our purpose. From Assumption 6 it is easy to see that

$$\|F(x) - F(z) - F'(z)(x-z)\| \leq \frac{1}{2}(K_1 + K_2)\|x-z\|\|F'(z)(x-z)\| \tag{2.18}$$

and

$$\|F(x)-F(z)-F'(z)(x-z)\| \le \frac{3}{2}(K_1+K_2)\|x-z\|\|F'(x)(x-z)\|. \tag{2.19}$$

for any $x,z \in B_\rho(x^\dagger)$.

We still need to deal with some commutators involving r_α. The structure information on F will be incorporated into such estimates. Thus, instead of Assumption 4, we need the following strengthened version.

Assumption 7 *(a) Under Assumption 5, there exists a positive constant c_7 such that*

$$\left\|r_\alpha\left(F'(x)^*F'(x)\right)-r_\alpha\left(F'(z)^*F'(z)\right)\right\| \le c_7K_0\|x-z\| \tag{2.20}$$

for all $x,z \in B_\rho(x^\dagger)$ and all $\alpha > 0$.

(b) Under Assumption 5 and Assumption 6, there exists a positive constant c_8 such that

$$\begin{aligned}&\|F'(x)\left[r_\alpha\left(F'(x)^*F'(x)\right)-r_\alpha\left(F'(z)^*F'(z)\right)\right]\|\\ &\qquad \le c_8(K_0+K_1)\alpha^{1/2}\|x-z\|+c_8K_2\left(\|F'(x)(x-z)\|+\|F'(z)(x-z)\|\right)\end{aligned} \tag{2.21}$$

for all $x,z \in B_\rho(x^\dagger)$ and all $\alpha > 0$.

Now we are ready to state the second main result in this paper which in particular says that the method (1.3) together with the discrepancy principle (1.7) defines an order optimal regularization method for each $0<\nu\le\bar{\nu}-1/2$ under stronger conditions on F. We will fix a constant $\gamma_1 > c_3r^{1/2}/(\tau-1)$.

Theorem 2 *Let $\{g_\alpha\}$ and $\{\alpha_k\}$ satisfy Assumption 1, (1.4), Assumption 2 and Assumption 7, let $\bar{\nu}\ge 1$ be the qualification of the linear regularization method defined by $\{g_\alpha\}$, and let F satisfy (2.8), (2.9), Assumption 5 and Assumption 6 with $\rho > 2(1+c_4\gamma_1)\|x_0-x^\dagger\|$. Let $\{x_k^\delta\}$ be defined by (1.3) and let k_δ be the first integer satisfying (1.7) with $\tau > 1$. Then there exists a constant $\eta_1 > 0$ depending only on r, τ and c_i, $i=0,\cdots,8$, such that if $(K_0+K_1+K_2)\|x_0-x^\dagger\| \le \eta_1$ then*

(i) If $x_0-x^\dagger$ satisfies the Hölder source condition (1.10) for some $\omega\in X$ and $0<\nu\le\bar{\nu}-1/2$, then

$$\|x_{k_\delta}^\delta-x^\dagger\| \le C_\nu\|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}, \tag{2.22}$$

where C_ν is a constant depending only on r, τ, ν and c_i, $i=0,\cdots,8$.

(ii) If $x_0-x^\dagger$ satisfies the logarithmic source condition (1.11) for some $\omega\in X$ and $\mu>0$, then

$$\|x_{k_\delta}^\delta-x^\dagger\| \le C_\mu\|\omega\|\left(1+\left|\ln\frac{\delta}{\|\omega\|}\right|\right)^{-\mu}, \tag{2.23}$$

where C_μ is a constant depending only on r, τ, μ, and c_i, $i=0,\cdots,8$.

In the statements of Theorem 1 and Theorem 2, the smallness of $L\|u\|$ and $(K_0+K_1+K_2)\|x_0-x^\dagger\|$ are not specified. However, during the proof of Theorem 1, we indeed will spell out all the necessary smallness conditions on $L\|u\|$. For simplicity of presentation, we will not spell out the smallness conditions on $(K_0+K_1+K_2)\|x_0-x^\dagger\|$ any more; the readers should be able to figure out such conditions without any difficulty.

Note that, without any source condition on $x_0-x^\dagger$, the above two theorems do not give the convergence of $x_{k_\delta}^\delta$ to $x^\dagger$. The following theorem says that $x_{k_\delta}^\delta \to x^\dagger$ as $\delta\to 0$ provided $x_0-x^\dagger\in\mathscr{N}(F'(x^\dagger))^\perp$. In fact, it tells more, it says that the convergence rates can even be improved to $o(\delta^{2\nu/(1+2\nu)})$ if $x_0-x^\dagger$ satisfies (1.10) for $0\le\nu<\bar{\nu}-1/2$.

Theorem 3 *(i) Let all the conditions in Theorem 1 be fulfilled. If $\bar{\nu} > 1$ and $x^\dagger - x_0$ satisfies the Hölder source condition (1.10) for some $\omega \in \mathscr{N}(F'(x^\dagger))^\perp$ and $1/2 \le \nu < \bar{\nu} - 1/2$, then*

$$\|x_{k_\delta}^\delta - x^\dagger\| \le o(\delta^{2\nu/(1+2\nu)})$$

as $\delta \to 0$.

(ii) Let all the conditions in Theorem 2 be fulfilled. If $x_0 - x^\dagger$ satisfies (1.10) for some $\omega \in \mathscr{N}(F'(x^\dagger))^\perp$ and $0 \le \nu < \bar{\nu} - 1/2$, then

$$\|x_{k_\delta}^\delta - x^\dagger\| \le o(\delta^{2\nu/(1+2\nu)})$$

as $\delta \to 0$.

Theorem 1, Theorem 2 and Theorem 3 will be proved in Sections 3, 4 and 5 respectively. In the following we will give some remarks.

Remark 1 A comprehensive overview on iterative regularization methods for nonlinear ill-posed problems may be found in the recent book [14]. In particular, convergence and convergence rates for the general method (1.3) are obtained in [14, Theorem 4.16] in case of a priori stopping rules under suitable nonlinearity assumptions on F.

Remark 2 In [18] Tautenhahn introduced a general regularization scheme for (1.1) by defining the regularized solutions x_α^δ as a fixed point of the nonlinear equation

$$x = x_0 - g_\alpha\left(F'(x)^*F'(x)\right)F'(x)^*\left(F(x) - y^\delta - F'(x)(x - x_0)\right), \tag{2.24}$$

where $\alpha > 0$ is the regularization parameter. When α is determined by a Morozov's type discrepancy principle, it was shown in [18] that the method is order optimal for each $0 < \nu \le \bar{\nu}/2$ under certain conditions on F. We point out that the technique developed in the present paper can be used to analyze such method; indeed we can even show that, under merely the Lipschitz condition on F', the method in [18] is order optimal for each $1/2 \le \nu \le \bar{\nu} - 1/2$, which improves the corresponding result.

Remark 3 Alternative to (1.3), one may consider the inexact Newton type methods

$$x_{k+1}^\delta = x_k^\delta - g_{\alpha_k}\left(F'(x_k^\delta)^*F'(x_k^\delta)\right)F'(x_k^\delta)^*\left(F(x_k^\delta) - y^\delta\right) \tag{2.25}$$

which can be derived by applying the regularization method defined by $\{g_\alpha\}$ to (1.5) with the current iterate x_k^δ as an initial guess. Such methods have first been studied by Hanke in [5,6] where the regularization properties of the Levenberg-Marquardt algorithm and the Newton-CG algorithm have been established without giving convergence rates when the sequence $\{\alpha_k\}$ is chosen adaptively during computation and the discrepancy principle is used as a stopping rule. The general methods (2.25) have been considered later by Rieder in [15,16], where $\{\alpha_k\}$ is determined by a somewhat different adaptive strategy; certain sub-optimal convergence rates have been derived when $x_0 - x^\dagger$ satisfies (1.10) with $\eta < \nu \le 1/2$ for some problem-dependent number $0 < \eta < 1/2$, while it is not yet clear if the convergence can be established under weaker source conditions. The convergence analysis of (2.25) is indeed far from complete. The technique in the present paper does not work for such methods.

Throughout this paper we will use $\{x_k\}$ to denote the iterated solutions defined by (1.3) corresponding to the noise free case. i.e.

$$x_{k+1} = x_0 - g_{\alpha_k}\left(F'(x_k)^*F'(x_k)\right)F'(x_k)^*\left(F(x_k) - y - F'(x_k)(x_k - x_0)\right). \tag{2.26}$$

We will also use the notations

$$\begin{aligned}
&\mathscr{A} := F'(x^\dagger)^*F'(x^\dagger), \quad \mathscr{A}_k := F'(x_k)^*F'(x_k), \quad \mathscr{A}_k^\delta := F'(x_k^\delta)^*F'(x_k^\delta),\\
&\mathscr{B} := F'(x^\dagger)F'(x^\dagger)^*, \quad \mathscr{B}_k := F'(x_k)F'(x_k)^*, \quad \mathscr{B}_k^\delta := F'(x_k^\delta)F'(x_k^\delta)^*,
\end{aligned}$$

and

$$e_k := x_k - x^\dagger, \qquad e_k^\delta := x_k^\delta - x^\dagger.$$

For ease of exposition, we will use C to denote a generic constant depending only on r. τ and c_i, $i = 0, \cdots, 8$, we will also use the convention $\Phi \lesssim \Psi$ to mean that $\Phi \le C\Psi$ for some generic constant C. Moreover, when we say $L\|u\|$ (or $(K_0+K_1+K_2)\|e_0\|$) is sufficiently small we will mean that $L\|u\| \le \eta$ (or $(K_0+K_1+K_2)\|e_0\| \le \eta$) for some small positive constant η depending only on r, τ and c_i, $i = 0, \cdots, 8$.

3 Proof of Theorem 1

In this section we will give the proof of Theorem 1. The main idea behind the proof consists of the following steps:

• Show the method defined by (1.3) and (1.7) is well-defined.
• Establish the stability estimate $\|x_k^\delta - x_k\| \lesssim \delta/\sqrt{\alpha_k}$. This enables us to write $\|e_{k_\delta}^\delta\| \lesssim \|e_{k_\delta}\| + \delta/\sqrt{\alpha_{k_\delta}}$.
• Establish $\alpha_{k_\delta} \ge C_\nu(\delta/\|\omega\|)^{2/(1+2\nu)}$ under the source condition (1.10) for $1/2 \le \nu \le \bar{\nu} - 1/2$. This is an easy step although it requires nontrivial arguments.
• Show $\|e_{k_\delta}\| \le C_\nu \|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}$, which is the hard part in the whole proof. In order to achieve this, we pick an integer $\bar{k}_\delta$ such that $k_\delta \le \bar{k}_\delta$ and $\alpha_{\bar{k}_\delta} \sim (\delta/\|\omega\|)^{2/(1+2\nu)}$. Such $\bar{k}_\delta$ will be proved to exist. Then we connect $\|e_{k_\delta}\|$ and $\|e_{\bar{k}_\delta}\|$ by establishing the inequality

$$\|e_{k_\delta}\| \lesssim \|e_{\bar{k}_\delta}\| + \frac{1}{\sqrt{\alpha_{\bar{k}_\delta}}}\left(\|F(x_{k_\delta}) - y\| + \delta\right). \tag{3.1}$$

The right hand side can be easily estimated by the desired bound.
• In order to establish (3.1), we need to establish the preliminary convergence rate estimate $\|e_{k_\delta}^\delta\| \lesssim \|u\|^{1/2}\delta^{1/2}$ when $x_0 - x^\dagger = F'(x^\dagger)^* u$ for some $u \in \mathscr{N}(F'(x^\dagger)^*)^\perp \subset Y$.

Therefore, in order to complete the proof of Theorem 1, we need to establish various estimates.

3.1 A first result on convergence rates

In this subsection we will derive the convergence rate $\|e_{k_\delta}^\delta\| \lesssim \|u\|^{1/2}\delta^{1/2}$ under the source condition

$$x_0 - x^\dagger = F'(x^\dagger)^* u, \quad u \in \mathscr{N}(F'(x^\dagger)^*)^\perp. \tag{3.2}$$

To this end, we introduce $\tilde{k}_\delta$ to be the first integer such that

$$\alpha_{\tilde{k}_\delta} \le \frac{\delta}{\gamma_0\|u\|} < \alpha_k, \qquad 0 \le k < \tilde{k}_\delta, \tag{3.3}$$

where γ_0 is a number satisfying $\gamma_0 > c_0 r/(\tau - 1)$, and c_0 is the constant from Assumption 1 (a). Because of (1.4), such $\tilde{k}_\delta$ is well-defined.

Theorem 4 *Let $\{g_\alpha\}$ and $\{\alpha_k\}$ satisfy Assumption 1(a), Assumption 2, (2.12) and (1.4), and let F satisfy (2.8), (2.9) and Assumption 3 with $\rho > 4\|x_0 - x^\dagger\|$. Let $\{x_k^\delta\}$ be defined by (1.3) and let k_δ be determined by the discrepancy principle (1.7) with $\tau > 1$. If $x_0 - x^\dagger$ satisfies (3.2) and if $L\|u\|$ is sufficiently small, then*

(i) For all $0 \le k \le \tilde{k}_\delta$ *there hold*

$$x_k^\delta \in B_\rho(x^\dagger) \qquad \textit{and} \qquad \|e_k^\delta\| \le 2(c_3 + c_4\gamma_0) r^{1/2} \alpha_k^{1/2} \|u\|. \tag{3.4}$$

(ii) $k_\delta \le \tilde{k}_\delta$*, i.e. the discrepancy principle (1.7) is well-defined.*

(iii) There exists a generic constant $C > 0$ *such that*

$$\|e_{k_\delta}^\delta\| \le C\|u\|^{1/2}\delta^{1/2}.$$

Proof We first prove (i). Note that $\rho > 4\|x_0 - x^\dagger\|$, it follows from (3.2) and (2.9) that (3.4) is trivial for $k = 0$. Now for any fixed integer $0 < l \le \tilde{k}_\delta$, we assume that (3.4) is true for all $0 \le k < l$. It follows from the definition (1.3) of $\{x_k^\delta\}$ that

$$e_{k+1}^\delta = r_{\alpha_k}(\mathscr{A}_k^\delta) e_0 - g_{\alpha_k}(\mathscr{A}_k^\delta) F'(x_k^\delta)^* \left(F(x_k^\delta) - y^\delta - F'(x_k^\delta) e_k^\delta \right). \tag{3.5}$$

Using (3.2), Assumption 3, Assumption 1(a), (2.1) and (1.2) we obtain

$$\begin{aligned} \|e_{k+1}^\delta\| &\le \|r_{\alpha_k}(\mathscr{A}_k^\delta) F'(x_k^\delta)^* u\| + \|r_{\alpha_k}(\mathscr{A}_k^\delta)[F'(x^\dagger)^* - F'(x_k^\delta)^*]u\| \\ &\quad + c_4 \alpha_k^{-1/2} \|F(x_k^\delta) - y^\delta - F'(x_k^\delta) e_k^\delta\| \\ &\le c_3 \alpha_k^{1/2}\|u\| + L\|u\|\|e_k^\delta\| + \frac{1}{2} c_4 L \|e_k^\delta\|^2 \alpha_k^{-1/2} + c_4 \delta \alpha_k^{-1/2}. \end{aligned}$$

Note that $\delta \alpha_k^{-1} \le \gamma_0 \|u\|$ for $0 \le k < \tilde{k}_\delta$. Note also that $\alpha_k \le r\alpha_{k+1}$ by (1.4). Therefore, by using (3.4) with $k = l - 1$, we obtain

$$\begin{aligned} \|e_l^\delta\| &\le r^{1/2} \alpha_l^{1/2} \left[(c_3 + c_4\gamma_0)\|u\| + L\|u\| \frac{\|e_{l-1}^\delta\|}{\sqrt{\alpha_{l-1}}} + \frac{1}{2} c_4 L \left(\frac{\|e_{l-1}^\delta\|}{\sqrt{\alpha_{l-1}}} \right)^2 \right] \\ &\le 2(c_3 + c_4\gamma_0) r^{1/2} \alpha_l^{1/2} \|u\| \end{aligned}$$

if $L\|u\|$ is so small that

$$2\left(r^{1/2} + (c_3 + c_4\gamma_0) c_4 r \right) L\|u\| \le 1. \tag{3.6}$$

By using (3.5), (2.1), Assumption 3, (1.2), Assumption 1(a), (3.4) with $k = l - 1$ and (3.6), we also obtain

$$\begin{aligned} \|e_l^\delta\| &\le \|r_{\alpha_{l-1}}(\mathscr{A}_{l-1}^\delta) e_0\| + c_4 \delta \alpha_{l-1}^{-1/2} + \frac{1}{2} c_4 L \|e_{l-1}^\delta\|^2 \alpha_{l-1}^{-1/2} \\ &\le \|e_0\| + c_4 \gamma_0^{1/2} \|u\|^{1/2} \delta^{1/2} + (c_3 + c_4\gamma_0) c_4 r^{1/2} L\|u\| \|e_{l-1}^\delta\| \\ &\le \|e_0\| + c_4 \gamma_0^{1/2} \|u\|^{1/2} \delta^{1/2} + \frac{1}{2}\rho \end{aligned}$$

Therefore, by using $\rho > 4\|e_0\|$, we have

$$\|e_k^\delta\| \le \frac{3}{4}\rho + c_4 \gamma_0^{1/2} \|u\|^{1/2} \delta^{1/2} < \rho$$

if $\delta > 0$ is small enough. Thus (3.4) is also true for all $k = l$. As $l \le \tilde{k}_\delta$ has been arbitrary, we have completed the proof of (i).

Next we prove (ii) by showing that $k_\delta \le \tilde{k}_\delta$. From (3.5) and (3.2) we have for $0 \le k < \tilde{k}_\delta$ that

$$\begin{aligned} F'(x^\dagger) e_{k+1}^\delta - y^\delta + y &= F'(x_k^\delta) r_{\alpha_k}(\mathscr{A}_k^\delta) \left[F'(x_k^\delta)^* + \left(F'(x^\dagger)^* - F'(x_k^\delta)^* \right) \right] u \\ &\quad + \left[F'(x^\dagger) - F'(x_k^\delta) \right] r_{\alpha_k}(\mathscr{A}_k^\delta) \left[F'(x_k^\delta)^* + \left(F'(x^\dagger)^* - F'(x_k^\delta)^* \right) \right] u \\ &\quad - \left[F'(x^\dagger) - F'(x_k^\delta) \right] g_{\alpha_k}(\mathscr{A}_k^\delta) F'(x_k^\delta)^* \left[F(x_k^\delta) - y^\delta - F'(x_k^\delta) e_k^\delta \right] \\ &\quad - g_{\alpha_k}(\mathscr{B}_k^\delta) \mathscr{B}_k^\delta \left[F(x_k^\delta) - y - F'(x_k^\delta) e_k^\delta \right] \\ &\quad - r_{\alpha_k}(\mathscr{B}_k^\delta)(y^\delta - y). \end{aligned}$$

By using Assumption 3, Assumption 1(a), (2.1), (1.2) and (3.4), and noting that $\delta/\alpha_k \le \gamma_0\|u\|$, we obtain

$$\begin{aligned}\|F'(x^\dagger)e_{k+1}^\delta - y^\delta + y\| &\le \delta + c_0\alpha_k\|u\| + 2c_3L\|u\|\alpha_k^{1/2}\|e_k^\delta\| + L^2\|u\|\|e_k^\delta\|^2 \\ &\qquad + c_4L\|e_k^\delta\|\delta\alpha_k^{-1/2} + \frac{1}{2}c_4L^2\alpha_k^{-1/2}\|e_k^\delta\|^3 + \frac{1}{2}L\|e_k^\delta\|^2 \\ &\le \delta + (c_0+\varepsilon_1)\,\alpha_k\|u\|,\end{aligned}$$

where

$$\begin{aligned}\varepsilon_1 &= \left[2r^{1/2}(c_3+c_4\gamma_0)(2c_3+c_4\gamma_0) + 2(c_3+c_4\gamma_0)^2 r\right] L\|u\| \\ &\quad + 4\left[(c_3+c_4\gamma_0)^2 r + (c_3+c_4\gamma_0)^3 c_4 r^{3/2}\right] L^2\|u\|^2.\end{aligned}$$

From (1.2), (3.2) and (2.9) we have $\|F'(x^\dagger)e_0 - y^\delta + y\| \le \delta + \|\mathscr{A}u\| \le \delta + c_0\alpha_0\|u\|$. Thus, by using (1.4),

$$\|F'(x^\dagger)e_k^\delta - y^\delta + y\| \le \delta + r\,(c_0+\varepsilon_1)\,\alpha_k\|u\|, \quad 0 \le k \le \tilde{k}_\delta.$$

Consequently

$$\begin{aligned}\|F(x_{\tilde{k}_\delta}^\delta) - y^\delta\| &\le \|F'(x^\dagger)e_{\tilde{k}_\delta}^\delta - y^\delta + y\| + \|F(x_{\tilde{k}_\delta}^\delta) - y - F'(x^\dagger)e_{\tilde{k}_\delta}^\delta\| \\ &\le \delta + r\,(c_0+\varepsilon_1)\,\alpha_{\tilde{k}_\delta}\|u\| + \frac{1}{2}L\|e_{\tilde{k}_\delta}^\delta\|^2 \\ &\le \delta + r\left(c_0+\varepsilon_1 + 2(c_3+c_4\gamma_0)^2 rL\|u\|\right)\alpha_{\tilde{k}_\delta}\|u\| \\ &\le \delta + r\left(c_0+\varepsilon_1 + 2(c_3+c_4\gamma_0)^2 rL\|u\|\right)\gamma_0^{-1}\delta \\ &\le \tau\delta\end{aligned}$$

if $L\|u\|$ is so small that

$$\varepsilon_1 + 2(c_3+c_4\gamma_0)^2 rL\|u\| \le \frac{(\tau-1)\gamma_0 - c_0 r}{r}.$$

By the definition of k_δ, it follows that $k_\delta \le \tilde{k}_\delta$.

Finally we are in a position to derive the convergence rate in (iii). If $k_\delta = 0$, then, by the definition of k_δ, we have $\|F(x_0) - y^\delta\| \le \tau\delta$. This together with Assumption 3 and (1.2) gives

$$\|F'(x^\dagger)e_0\| \le \|F(x_0) - y - F'(x^\dagger)e_0\| + \|F(x_0) - y\| \le \frac{1}{2}L\|e_0\|^2 + (\tau+1)\delta.$$

Thus, by using (3.2), we have

$$\begin{aligned}\|e_0\| &= (e_0, F'(x^\dagger)^* u)^{1/2} = (F'(x^\dagger)e_0, u)^{1/2} \le \|F'(x^\dagger)e_0\|^{1/2}\|u\|^{1/2} \\ &\le \sqrt{\frac{1}{2}L\|u\|}\|e_0\| + \sqrt{\tau+1}\|u\|^{1/2}\delta^{1/2}.\end{aligned}$$

By assuming that $L\|u\| \le 1$, we obtain $\|e_{k_\delta}^\delta\| = \|e_0\| \lesssim \|u\|^{1/2}\delta^{1/2}$.

Therefore we will assume $k_\delta > 0$ in the following argument. It follows from (3.5), (2.1), Assumption 3 and (3.4) that for $0 \le k < \tilde{k}_\delta$

$$\begin{aligned}\|e_{k+1}^\delta\| &\le \|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\| + c_4\delta\alpha_k^{-1/2} + \frac{1}{2}c_4L\|e_k^\delta\|^2\alpha_k^{-1/2} \\ &\le \|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\| + c_4(\gamma_0\|u\|\delta)^{1/2} + (c_3+c_4\gamma_0)c_4r^{1/2}L\|u\|\|e_k^\delta\|. \qquad (3.7)\end{aligned}$$

By (3.2), (2.12) in Assumption 4, and Assumption 3 we have

$$\begin{aligned} \|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0 - r_{\alpha_k}(\mathscr{A})e_0\| &= \|[r_{\alpha_k}(\mathscr{A}_k^\delta) - r_{\alpha_k}(\mathscr{A})]F'(x^\dagger)^* u\| \\ &\le c_6\|u\|\|F'(x_k^\delta) - F'(x^\dagger)\| \\ &\le c_6 L\|u\|\|e_k^\delta\|. \end{aligned} \tag{3.8}$$

Thus

$$\begin{aligned} \|e_{k+1}^\delta\| &\le \|r_{\alpha_k}(\mathscr{A})e_0\| + c_4(\gamma_0\|u\|\delta)^{1/2} + \left(c_6 + (c_3 + c_4\gamma_0)c_4 r^{1/2}\right) L\|u\|\|e_k^\delta\| \\ &\le \|r_{\alpha_k}(\mathscr{A})e_0\| + c_4(\gamma_0\|u\|\delta)^{1/2} + \frac{1}{4c_5}\|e_k^\delta\| \end{aligned} \tag{3.9}$$

if we assume further that

$$4c_5\left(c_6 + (c_3 + c_4\gamma_0)c_4 r^{1/2}\right) L\|u\| \le 1. \tag{3.10}$$

Note that (2.9) and the choice of β_0 imply $\|r_{\alpha_0}(\mathscr{A})e_0\| \ge \frac{3}{4}\|e_0\|$. Thus, with the help of (2.7), by induction we can conclude from (3.9) that

$$\|e_k^\delta\| \le \frac{4}{3}c_5\|r_{\alpha_k}(\mathscr{A})e_0\| + C\|u\|^{1/2}\delta^{1/2}, \quad 0 \le k \le \tilde{k}_\delta.$$

This together with (3.8) and (3.10) implies

$$\|e_k^\delta\| \le 2c_5\|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\| + C\|u\|^{1/2}\delta^{1/2}, \quad 0 \le k \le \tilde{k}_\delta. \tag{3.11}$$

The combination of (3.7), (3.11) and (3.10) gives

$$\|e_{k+1}^\delta\| \le \frac{3}{2}\|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\| + C\|u\|^{1/2}\delta^{1/2}, \quad 0 \le k < \tilde{k}_\delta. \tag{3.12}$$

We need to estimate $\|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\|$. By (3.2), Assumption 1(a) and Assumption 3 we have

$$\begin{aligned} \|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\|^2 &= \left(r_{\alpha_k}(\mathscr{A}_k^\delta)e_0, r_{\alpha_k}(\mathscr{A}_k^\delta)F'(x^\dagger)^* u\right) \\ &= \left(r_{\alpha_k}(\mathscr{A}_k^\delta)e_0, r_{\alpha_k}(\mathscr{A}_k^\delta)\left[F'(x_k^\delta)^* + \left(F'(x^\dagger)^* - F'(x_k^\delta)^*\right)\right] u\right) \\ &\le \|F'(x_k^\delta)r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\|\|u\| + L\|u\|\|e_k^\delta\|\|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\|. \end{aligned}$$

Thus

$$\|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\| \le \|F'(x_k^\delta)r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\|^{1/2}\|u\|^{1/2} + L\|u\|\|e_k^\delta\|.$$

With the help of (3.5), (1.2), Assumption 1(a) and Assumption 3 we have

$$\begin{aligned} \|F'(x_k^\delta)r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\| &\le \|F'(x_k^\delta)e_{k+1}^\delta\| + \|g_{\alpha_k}(\mathscr{B}_k^\delta)\mathscr{B}_k^\delta\left(F(x_k^\delta) - y^\delta - F'(x_k^\delta)e_k^\delta\right)\| \\ &\le \|F(x_{k+1}^\delta) - y^\delta\| + 2\delta + \|F(x_{k+1}^\delta) - y - F'(x_{k+1}^\delta)e_{k+1}^\delta\| \\ &\quad + \|[F'(x_{k+1}^\delta) - F'(x_k^\delta)]e_{k+1}^\delta\| + \|F(x_k^\delta) - y - F'(x_k^\delta)e_k^\delta\| \\ &\le \|F(x_{k+1}^\delta) - y^\delta\| + 2\delta + L\|e_k^\delta\|^2 + 2L\|e_{k+1}^\delta\|^2. \end{aligned}$$

Therefore

$$\begin{aligned} \|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\| &\le \|u\|^{1/2}\|F(x_{k+1}^\delta) - y^\delta\|^{1/2} + \sqrt{2}\|u\|^{1/2}\delta^{1/2} + \sqrt{2L\|u\|}\|e_{k+1}^\delta\| \\ &\quad + \left(L\|u\| + \sqrt{L\|u\|}\right)\|e_k^\delta\|. \end{aligned}$$

Combining this with (3.11) and (3.12) yields

$$\|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\| \le \|u\|^{1/2}\|F(x_{k+1}^\delta)-y^\delta\|^{1/2}+C\|u\|^{1/2}\delta^{1/2} + \frac{1}{2}\left[\left(3\sqrt{2}+4c_5\right)\sqrt{L\|u\|}+4c_5L\|u\|\right]\|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\|.$$

Thus, if

$$\left(3\sqrt{2}+4c_5\right)\sqrt{L\|u\|}+4c_5L\|u\|\le 1,$$

we then obtain

$$\|r_{\alpha_k}(\mathscr{A}_k^\delta)e_0\| \lesssim \|u\|^{1/2}\|F(x_{k+1}^\delta)-y^\delta\|^{1/2}+\|u\|^{1/2}\delta^{1/2}.$$

This together with (3.12) gives

$$\|e_k^\delta\| \lesssim \|u\|^{1/2}\|F(x_k^\delta)-y^\delta\|^{1/2}+\|u\|^{1/2}\delta^{1/2}$$

for all $0<k\le\tilde{k}_\delta$. Consequently, we may set $k=k_\delta$ in the above inequality and use the definition of k_δ to obtain $\|e_{k_\delta}^\delta\|\lesssim\|u\|^{1/2}\delta^{1/2}$. □

3.2 **Stability estimates**

In this subsection we will consider the stability of the method (1.3) by deriving some useful estimates on $\|x_k^\delta-x_k\|$, where $\{x_k\}$ is defined by (2.26). It is easy to see that

$$e_{k+1}=r_{\alpha_k}(\mathscr{A}_k)e_0-g_{\alpha_k}(\mathscr{A}_k)F'(x_k)^*\left(F(x_k)-y-F'(x_k)e_k\right). \tag{3.13}$$

We will prove some important estimates on $\{x_k\}$ in Lemma 3 in the next subsection. In particular, we will show that, under the conditions in Theorem 4,

$$x_k\in B_\rho(x^\dagger) \qquad \text{and} \qquad \|e_k\|\le 2c_3r^{1/2}\alpha_k^{1/2}\|u\| \tag{3.14}$$

for all $k\ge 0$ provided $L\|u\|$ is sufficiently small.

Lemma 2 *Let all the conditions in Theorem 4 and Assumption 4 hold. If $L\|u\|$ is sufficiently small, then for all $0\le k\le\tilde{k}_\delta$ there hold*

$$\|x_k^\delta-x_k\|\le 2c_4\frac{\delta}{\sqrt{\alpha_k}} \tag{3.15}$$

and

$$\|F(x_k^\delta)-F(x_k)-y^\delta+y\|\le(1+\varepsilon_2)\delta, \tag{3.16}$$

where

$$\varepsilon_2:=2c_4\left((c_6+rc_4\gamma_0)+(4c_3+3c_4\gamma_0)r^{1/2}+4(c_3+c_4\gamma_0)r\right)L\|u\| + 4c_3c_4\left(c_6r^{1/2}+(c_4+c_6)c_3r\right)L^2\|u\|^2.$$

Proof For each $0\le k\le\tilde{k}_\delta$ we set

$$u_k:=F(x_k)-y-F'(x_k)e_k, \qquad u_k^\delta:=F(x_k^\delta)-y-F'(x_k^\delta)e_k^\delta. \tag{3.17}$$

It then follows from (3.5) and (3.13) that

$$x_{k+1}^\delta-x_{k+1}=I_1+I_2+I_3+I_4, \tag{3.18}$$

where

$$
\begin{aligned}
I_1 &:= \left[r_{\alpha_k}(\mathscr{A}_k^\delta) - r_{\alpha_k}(\mathscr{A}_k)\right] e_0,\\
I_2 &:= g_{\alpha_k}(\mathscr{A}_k^\delta) F'(x_k^\delta)^*(y^\delta - y),\\
I_3 &:= \left[g_{\alpha_k}(\mathscr{A}_k) F'(x_k)^* - g_{\alpha_k}(\mathscr{A}_k^\delta) F'(x_k^\delta)^*\right] u_k,\\
I_4 &:= g_{\alpha_k}(\mathscr{A}_k^\delta) F'(x_k^\delta)^*(u_k - u_k^\delta).
\end{aligned}
$$

By using (3.2), (2.11), (2.12), Assumption 3 and (3.14) we have

$$
\begin{aligned}
\|I_1\| &\le \|r_{\alpha_k}(\mathscr{A}_k^\delta) - r_{\alpha_k}(\mathscr{A}_k)\| \|F'(x^\dagger)^* - F'(x_k)^*\| \|u\|\\
&\quad + \|[r_{\alpha_k}(\mathscr{A}_k^\delta) - r_{\alpha_k}(\mathscr{A}_k)] F'(x_k)^* u\|\\
&\le c_6 L^2 \|u\| \|e_k\| \|x_k^\delta - x_k\| \alpha_k^{-1/2} + c_6 L \|u\| \|x_k^\delta - x_k\|\\
&\le c_6 \left(L\|u\| + 2c_3 r^{1/2} L^2 \|u\|^2\right) \|x_k^\delta - x_k\|.
\end{aligned}
$$

With the help of (2.1) and (1.2) we have

$$
\|I_2\| \le c_4 \frac{\delta}{\sqrt{\alpha_k}}.
$$

By applying Assumption 1(a), (2.14), Assumption 3 and (3.14) we can estimate I_3 as

$$
\begin{aligned}
\|I_3\| &\le \|g_{\alpha_k}(\mathscr{A}_k)[F'(x_k^\delta)^* - F'(x_k)^*] u_k\| + \|[g_{\alpha_k}(\mathscr{A}_k) - g_{\alpha_k}(\mathscr{A}_k^\delta)] F'(x_k^\delta)^* u_k\|\\
&\le (c_1 + c_6) L \|u_k\| \|x_k^\delta - x_k\| \alpha_k^{-1} \le \frac{1}{2}(c_1 + c_6) L^2 \|e_k\|^2 \|x_k^\delta - x_k\| \alpha_k^{-1}\\
&\le 2(c_1 + c_6) c_3^2 r L^2 \|u\|^2 \|x_k^\delta - x_k\|.
\end{aligned}
$$

For the term I_4, we have from (2.1) that

$$
\|I_4\| \le \frac{c_4}{\sqrt{\alpha_k}} \|u_k^\delta - u_k\|.
$$

By using Assumption 3, (3.4) and (3.14) one can see

$$
\begin{aligned}
\|u_k - u_k^\delta\| &\le \|F(x_k^\delta) - F(x_k) - F'(x_k)(x_k^\delta - x_k)\| + \|[F'(x_k^\delta) - F'(x_k)] e_k^\delta\|\\
&\le \frac{1}{2} L \|x_k^\delta - x_k\|^2 + L \|e_k^\delta\| \|x_k^\delta - x_k\| \le \frac{1}{2} L \left(3\|e_k^\delta\| + \|e_k\|\right) \|x_k^\delta - x_k\|\\
&\le (4c_3 + 3c_4\gamma_0) r^{1/2} \alpha_k^{1/2} L \|u\| \|x_k^\delta - x_k\|.
\end{aligned} \tag{3.19}
$$

Therefore

$$
\|I_4\| \le (4c_3 + 3c_4\gamma_0) c_4 r^{1/2} L \|u\| \|x_k^\delta - x_k\|.
$$

Thus, if $L\|u\|$ is so small that

$$
\left(c_6 + (4c_3 + 3c_4\gamma_0) c_4 r^{1/2}\right) L\|u\| + 2\left(c_3 c_6 r^{1/2} + c_3^2 (c_1 + c_6) r\right) L^2 \|u\|^2 \le \frac{1}{2},
$$

then the combination of the above estimates on I_1, I_2, I_3 and I_4 gives for $0 \le k < \tilde{k}_\delta$ that

$$
\|x_{k+1}^\delta - x_{k+1}\| \le c_4 \frac{\delta}{\sqrt{\alpha_k}} + \frac{1}{2} \|x_k^\delta - x_k\|.
$$

This implies (3.15) immediately.

Next we prove (3.16). We have from (3.18) that

$$
F'(x_k^\delta)(x_{k+1}^\delta - x_{k+1}) - y^\delta + y = F'(x_k^\delta)\left(I_1 + I_2 + I_3 + I_4\right) - y^\delta + y. \tag{3.20}
$$

From (3.2), (2.12), (2.13), Assumption 3, (3.14) and (3.15) it follows that

$$\begin{aligned}
\|F'(x_k^\delta)I_1\| &\le \|F'(x_k^\delta)[r_{\alpha_k}(\mathscr{A}_k^\delta)-r_{\alpha_k}(\mathscr{A}_k)][F'(x^\dagger)^*-F'(x_k)^*]u\| \\
&\quad + \|F'(x_k^\delta)[r_{\alpha_k}(\mathscr{A}_k^\delta)-r_{\alpha_k}(\mathscr{A}_k)]F'(x_k)^*u\| \\
&\le c_6L^2\|u\|\|e_k\|\|x_k^\delta-x_k\|+c_6L\|u\|\alpha_k^{1/2}\|x_k^\delta-x_k\| \\
&\le \left(2c_4c_6L\|u\|+4c_3c_4c_6r^{1/2}L^2\|u\|^2\right)\delta.
\end{aligned}$$

By using Assumption 1(a) and (1.2) it is easy to see

$$\|F'(x_k^\delta)I_2-y^\delta+y\|=\|r_{\alpha_k}(\mathscr{B}_k^\delta)(y^\delta-y)\|\le\delta. \tag{3.21}$$

In order to estimate $F'(x_k^\delta)I_3$, we note that

$$F'(x_k^\delta)I_3=\left[F'(x_k^\delta)-F'(x_k)\right]g_{\alpha_k}(\mathscr{A}_k)F'(x_k)^*u_k+\left[r_{\alpha_k}(\mathscr{B}_k^\delta)-r_{\alpha_k}(\mathscr{B}_k)\right]u_k. \tag{3.22}$$

Thus, it follows from (2.1), Assumption 3, (2.11), (3.14) and (3.15) that

$$\begin{aligned}
\|F'(x_k^\delta)I_3\| &\le \|\left[F'(x_k^\delta)-F'(x_k)\right]g_{\alpha_k}(\mathscr{A}_k)F'(x_k)^*u_k\| \\
&\quad + \|\left[r_{\alpha_k}(\mathscr{B}_k^\delta)-r_{\alpha_k}(\mathscr{B}_k)\right]u_k\| \\
&\le (c_4+c_6)\alpha_k^{-1/2}L\|x_k^\delta-x_k\|\|u_k\| \\
&\le \frac{1}{2}(c_4+c_6)\alpha_k^{-1/2}L^2\|e_k\|^2\|x_k^\delta-x_k\| \\
&\le 4(c_4+c_6)c_3^2c_4rL^2\|u\|^2\delta.
\end{aligned}$$

For the term $F'(x_k^\delta)I_4$ we have from Assumption 1(a), (3.19) and (3.15) that

$$\|F'(x_k^\delta)I_4\|\le\|u_k-u_k^\delta\|\le 2(4c_3+3c_4\gamma_0)c_4r^{1/2}L\|u\|\delta.$$

Combining the above estimates, we therefore obtain

$$\|F'(x_k^\delta)(x_{k+1}^\delta-x_{k+1})-y^\delta+y\|\le(1+\varepsilon_3)\delta, \qquad 0\le k<\tilde{k}_\delta,$$

where

$$\varepsilon_3 := 2c_4\left(c_6+(4c_3+3c_4\gamma_0)r^{1/2}\right)L\|u\|+4c_3c_4\left(c_6r^{1/2}+(c_4+c_6)c_3r\right)L^2\|u\|^2.$$

This together with Assumption 3, (3.4), (3.15) and (1.4) implies for $0\le k<\tilde{k}_\delta$ that

$$\begin{aligned}
&\|F'(x_{k+1}^\delta)(x_{k+1}^\delta-x_{k+1})-y^\delta+y\| \\
&\qquad \le \|F'(x_k^\delta)(x_{k+1}^\delta-x_{k+1})-y^\delta+y\|+L\|x_{k+1}^\delta-x_k^\delta\|\|x_{k+1}^\delta-x_{k+1}\| \\
&\qquad \le (1+\varepsilon_3)\delta+2c_4L(\|e_{k+1}^\delta\|+\|e_k^\delta\|)\frac{\delta}{\sqrt{\alpha_{k+1}}} \\
&\qquad \le (1+\varepsilon_4)\delta,
\end{aligned}$$

where

$$\varepsilon_4 := \varepsilon_3+8(c_3+c_4\gamma_0)c_4rL\|u\|.$$

Thus

$$\|F'(x_k^\delta)(x_k^\delta-x_k)-y^\delta+y\|\le(1+\varepsilon_4)\delta, \qquad 0\le k\le\tilde{k}_\delta.$$

Therefore, noting that $\delta/\alpha_k \le r\gamma_0\|u\|$ for $0 \le k \le \tilde{k}_\delta$, we have

$$\begin{aligned}
\|F(x_k^\delta) - F(x_k) - y^\delta + y\| &\le \|F(x_k^\delta) - F(x_k) - F'(x_k^\delta)(x_k^\delta - x_k)\| \\
&\quad + \|F'(x_k^\delta)(x_k^\delta - x_k) - y^\delta + y\| \\
&\le \frac{1}{2}L\|x_k^\delta - x_k\|^2 + (1+\varepsilon_4)\delta \\
&\le 2c_4^2 L \frac{\delta}{\alpha_k}\delta + (1+\varepsilon_4)\delta \\
&\le (1+\varepsilon_4 + 2rc_4^2\gamma_0 L\|u\|)\delta.
\end{aligned}$$

The proof of (3.16) is thus complete. □

3.3 **Some estimates on noise-free iterations**

Lemma 3 *Let all the conditions in Theorem 4 be fulfilled. If $L\|u\|$ is sufficiently small, then for all $k \ge 0$ we have*

$$x_k \in B_\rho(x^\dagger) \qquad \textit{and} \qquad \|e_k\| \le 2c_3 r^{1/2}\alpha_k^{1/2}\|u\|. \tag{3.23}$$

If, in addition, Assumption 1(b) is satisfied, then

$$\frac{2}{3}\|r_{\alpha_k}(\mathscr{A})e_0\| \le \|e_k\| \le \frac{4}{3}c_5\|r_{\alpha_k}(\mathscr{A})e_0\| \tag{3.24}$$

and

$$\frac{1}{2c_5}\|e_k\| \le \|e_{k+1}\| \le 2\|e_k\|. \tag{3.25}$$

Proof By using (3.2), (2.1), (2.12) and Assumption 3, we have from (3.13) that

$$\begin{aligned}
\|e_{k+1} - r_{\alpha_k}(\mathscr{A})e_0\| &\le \|[r_{\alpha_k}(\mathscr{A}_k) - r_{\alpha_k}(\mathscr{A})]F'(x^\dagger)^* u\| + \frac{c_4}{\sqrt{\alpha_k}}\|F(x_k) - y - F'(x_k)e_k\| \\
&\le c_6 L\|u\|\|e_k\| + \frac{c_4}{2\sqrt{\alpha_k}}L\|e_k\|^2.
\end{aligned} \tag{3.26}$$

Since (2.1) and (3.2) imply $\|r_{\alpha_k}(\mathscr{A})e_0\| \le c_3\alpha_k^{1/2}\|u\|$, we have

$$\|e_{k+1}\| \le c_3\alpha_k^{1/2}\|u\| + c_6 L\|u\|\|e_k\| + \frac{c_4}{2\sqrt{\alpha_k}}L\|e_k\|^2.$$

Note that (3.2) and (2.9) imply $\|e_0\| \le c_3\alpha_0^{1/2}\|u\|$. By induction one can conclude the assertion (3.23) if $L\|u\|$ is so small that $2(c_6 r^{1/2} + c_3 c_4 r)L\|u\| \le 1$.

If we assume further that

$$5c_5\left(c_6 + c_3c_4 r^{1/2}\right)L\|u\| \le 1, \tag{3.27}$$

the combination of (3.26) and (3.23) gives

$$\|e_{k+1} - r_{\alpha_k}(\mathscr{A})e_0\| \le \left(c_6 + c_3c_4r^{1/2}\right)L\|u\|\|e_k\| \le \frac{1}{5c_5}\|e_k\|. \tag{3.28}$$

Note that Assumption 1(b) and $\alpha_k \le \alpha_{k-1}$ imply $\|r_{\alpha_k}(\mathscr{A})e_0\| \le \|r_{\alpha_{k-1}}(\mathscr{A})e_0\|$. Note also that Assumption 1(a) and (2.9) imply (3.24) with $k=0$. Thus, from (3.28) and (2.7) we can conclude (3.24) by an induction argument. (3.25) is an immediate consequence of (3.28) and (3.24). □

Lemma 4 *Let all the conditions in Lemma 2 and Assumption 1(c) hold. If $k_\delta > 0$ and $L\|u\|$ is sufficiently small, then for all $k \ge k_\delta$ we have*

$$\|e_{k_\delta}\| \lesssim \|e_k\| + \frac{1}{\sqrt{\alpha_k}}\left(\|F(x_{k_\delta}) - y\| + \delta\right). \tag{3.29}$$

Proof It follows from (3.13) that

$$\begin{aligned} x_{k_\delta} - x_k = & [r_{\alpha_{k_\delta-1}}(\mathscr{A}) - r_{\alpha_{k-1}}(\mathscr{A})]e_0 + [r_{\alpha_{k_\delta-1}}(\mathscr{A}_{k_\delta-1}) - r_{\alpha_{k_\delta-1}}(\mathscr{A})]e_0 \\ & - \left[r_{\alpha_{k-1}}(\mathscr{A}_{k-1}) - r_{\alpha_{k-1}}(\mathscr{A})\right] e_0 \\ & - g_{\alpha_{k_\delta-1}}(\mathscr{A}_{k_\delta-1})F'(x_{k_\delta-1})^* \left[F(x_{k_\delta-1}) - y - F'(x_{k_\delta-1})e_{k_\delta-1}\right] \\ & + g_{\alpha_{k-1}}(\mathscr{A}_{k-1})F'(x_{k-1})^* \left[F(x_{k-1}) - y - F'(x_{k-1})e_{k-1}\right]. \end{aligned} \tag{3.30}$$

Thus, by using (3.2), (2.12), Assumption 3, (2.1), (3.23) and (3.27), we have

$$\begin{aligned} \|x_{k_\delta} - x_k\| \le & \|[r_{\alpha_{k_\delta-1}}(\mathscr{A}) - r_{\alpha_{k-1}}(\mathscr{A})]e_0\| + c_6 L\|u\| \left(\|e_{k-1}\| + \|e_{k_\delta-1}\|\right) \\ & + \frac{c_4}{2\sqrt{\alpha_{k_\delta-1}}} L\|e_{k_\delta-1}\|^2 + \frac{c_4}{2\sqrt{\alpha_{k-1}}} L\|e_{k-1}\|^2 \\ \le & \|[r_{\alpha_{k_\delta-1}}(\mathscr{A}) - r_{\alpha_{k-1}}(\mathscr{A})]e_0\| + \frac{1}{5c_5}\left(\|e_{k-1}\| + \|e_{k_\delta-1}\|\right). \end{aligned} \tag{3.31}$$

Since $k \ge k_\delta$, we have $\alpha_{k-1} \le \alpha_{k_\delta-1}$. Since Assumption 1(b) and (c) hold, we may apply Lemma 1 with $x = e_0$, $\bar{x} = e_{k_\delta}$, $\alpha = \alpha_{k-1}$, $\beta = \alpha_{k_\delta-1}$ and $A = F'(x^\dagger)$ to obtain

$$\|[r_{\alpha_{k_\delta-1}}(\mathscr{A}) - r_{\alpha_{k-1}}(\mathscr{A})]e_0\| \le \|r_{\alpha_{k_\delta-1}}(\mathscr{A})e_0 - e_{k_\delta}\| + \frac{c_2}{\sqrt{\alpha_{k-1}}}\|F'(x^\dagger)e_{k_\delta}\|.$$

Note that (3.28) implies

$$\|e_{k_\delta} - r_{\alpha_{k_\delta-1}}(\mathscr{A})e_0\| \le \frac{1}{5c_5}\|e_{k_\delta-1}\|.$$

Note also that Assumption 3 implies

$$\|F'(x^\dagger)e_{k_\delta}\| \le \|F(x_{k_\delta}) - y\| + \frac{1}{2}L\|e_{k_\delta}\|^2.$$

Thus

$$\|[r_{\alpha_{k_\delta-1}}(\mathscr{A}) - r_{\alpha_{k-1}}(\mathscr{A})]e_0\| \le \frac{1}{5c_5}\|e_{k_\delta-1}\| + \frac{C}{\sqrt{\alpha_k}}\left(\|F(x_{k_\delta}) - y\| + L\|e_{k_\delta}\|^2\right).$$

Since Lemma 2, Theorem 4 and the fact $k_\delta \le \tilde{k}_\delta$ imply

$$\|e_{k_\delta}\| \lesssim \|e^\delta_{k_\delta}\| + \frac{\delta}{\sqrt{\alpha_{k_\delta}}} \lesssim \|u\|^{1/2}\delta^{1/2},$$

we have

$$\|[r_{\alpha_{k_\delta-1}}(\mathscr{A}) - r_{\alpha_{k-1}}(\mathscr{A})]e_0\| \le \frac{1}{5c_5}\|e_{k_\delta-1}\| + \frac{C}{\sqrt{\alpha_k}}\left(\|F(x_{k_\delta}) - y\| + L\|u\|\delta\right).$$

Combining this with (3.31) and using Lemma 3 gives

$$\|x_{k_\delta} - x_k\| \le \frac{4}{5}\|e_{k_\delta}\| + C\|e_k\| + \frac{C}{\sqrt{\alpha_k}}\left(\|F(x_{k_\delta}) - y\| + \delta\right).$$

This completes the proof. □

3.4 **Completion of proof of Theorem 1**

Lemma 5 *Assume that all the conditions in Lemma 3 are satisfied. Then*

$$\|F'(x^\dagger)e_k\| \lesssim \|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + \alpha_k^{1/2}\|r_{\alpha_k}(\mathscr{A})e_0\| \tag{3.32}$$

for all $k \geq 0$.

Proof We first use (3.13) to write

$$\begin{aligned} F'(x^\dagger)e_{k+1} = F'(x^\dagger)r_{\alpha_k}(\mathscr{A})e_0 + F'(x^\dagger)\left[r_{\alpha_k}(\mathscr{A}_k) - r_{\alpha_k}(\mathscr{A})\right]e_0 \\ - F'(x^\dagger)g_{\alpha_k}(\mathscr{A}_k)F'(x_k)^*\left[F(x_k) - y - F'(x_k)e_k\right]. \end{aligned} \tag{3.33}$$

Thus, it follows from (3.2), Assumption 3, Assumption 1(a), (2.12), (2.13), (3.23) and (3.24) that

$$\begin{aligned} \|F'(x^\dagger)e_{k+1}\| &\lesssim \|F'(x^\dagger)r_{\alpha_k}(\mathscr{A})e_0\| + L\|e_k\|\|[r_{\alpha_k}(\mathscr{A}_k) - r_{\alpha_k}(\mathscr{A})]F'(x^\dagger)^*u\| \\ &\quad + \|F'(x_k)[r_{\alpha_k}(\mathscr{A}_k) - r_{\alpha_k}(\mathscr{A})]F'(x^\dagger)^*u\| + (1 + L\|e_k\|\alpha_k^{-1/2})L\|e_k\|^2 \\ &\lesssim \|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + L^2\|u\|\|e_k\|^2 + \alpha_k^{1/2}L\|u\|\|e_k\| + L\|e_k\|^2 \\ &\lesssim \|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + \alpha_k^{1/2}\|r_{\alpha_k}(\mathscr{A})e_0\|. \end{aligned}$$

This together with (2.7) and (1.4) implies (3.32). □

Lemma 6 *Under the conditions in Lemma 2 and Lemma 3, if* $\varepsilon_2 \leq (\tau - 1)/2$ *then for the* k_δ *determined by (1.7) with* $\tau > 1$ *we have*

$$(\tau - 1)\delta \lesssim \|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + \alpha_k^{1/2}\|r_{\alpha_k}(\mathscr{A})e_0\| \tag{3.34}$$

for all $0 \leq k < k_\delta$,

Proof By using (3.16), Lemma 3 and Lemma 5, we have for $0 \leq k < k_\delta$ that

$$\begin{aligned} \tau\delta &\leq \|F(x_k^\delta) - y^\delta\| \leq \|F(x_k^\delta) - F(x_k) - y^\delta + y\| + \|F(x_k) - y\| \\ &\leq (1 + \varepsilon_2)\delta + \|F'(x^\dagger)e_k\| + \frac{1}{2}L\|e_k\|^2 \\ &\leq (1 + \varepsilon_2)\delta + C\|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + C\alpha_k^{1/2}\|r_{\alpha_k}(\mathscr{A})e_0\|. \end{aligned}$$

Since $\tau > 1$, by the smallness condition $\varepsilon_2 \leq (\tau - 1)/2$ on $L\|u\|$ we obtain (3.34). □

Proof of Theorem 1. If $k_\delta = 0$, then the definition of k_δ implies $\|F(x_0) - y^\delta\| \leq \tau\delta$. From Theorem 4 we know that $\|e_0\| \lesssim \|u\|^{1/2}\delta^{1/2}$. Thus

$$\begin{aligned} \|F'(x^\dagger)e_0\| &\leq \|F(x_0) - y - F'(x^\dagger)e_0\| + \|F(x_0) - y^\delta\| + \delta \\ &\leq \frac{1}{2}L\|e_0\|^2 + (1 + \tau)\delta \lesssim \delta. \end{aligned}$$

Since $e_0 = \mathscr{A}^\nu\omega$ for some $1/2 \leq \nu \leq \bar{\nu} - 1/2$, we may use the interpolation inequality to obtain

$$\begin{aligned} \|e_{k_\delta}^\delta\| = \|e_0\| = \|\mathscr{A}^\nu\omega\| &\leq \|\omega\|^{1/(1+2\nu)}\|\mathscr{A}^{1/2+\nu}\omega\|^{2\nu/(1+2\nu)} \\ &= \|\omega\|^{1/(1+2\nu)}\|F'(x^\dagger)e_0\|^{2\nu/(1+2\nu)} \\ &\lesssim \|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}, \end{aligned}$$

which gives the desired estimate.

Therefore, we may assume that $k_\delta > 0$ in the remaining argument. By using $e_0 = \mathscr{A}^\nu\omega$ for some $1/2 \leq \nu \leq \bar{\nu} - 1/2$ and Lemma 6 it follows that there exists a positive constant C_ν such that

$$(\tau - 1)\delta < C_\nu\alpha_k^{\nu+1/2}\|\omega\|, \qquad 0 \leq k < k_\delta.$$

Now we define the integer $\bar{k}_\delta$ by

$$\alpha_{\bar{k}_\delta} \le \left(\frac{(\tau-1)\delta}{C_\nu\|\omega\|}\right)^{2/(1+2\nu)} < \alpha_k, \qquad 0 \le k < \bar{k}_\delta.$$

Then $k_\delta \le \bar{k}_\delta$. Thus, by using Lemma 2 and Lemma 4, we have

$$\|e_{k_\delta}^\delta\| \lesssim \|e_{k_\delta}\| + \frac{\delta}{\sqrt{\alpha_{k_\delta}}} \lesssim \|e_{\bar{k}_\delta}\| + \frac{\|F(x_{k_\delta})-y\|+\delta}{\sqrt{\alpha_{\bar{k}_\delta}}} + \frac{\delta}{\sqrt{\alpha_{k_\delta}}}.$$

Note that Lemma 2 and the definition of k_δ imply

$$\|F(x_{k_\delta})-y\| \le \|F(x_{k_\delta}^\delta)-y^\delta\| + \|F(x_{k_\delta}^\delta)-F(x_{k_\delta})-y^\delta+y\| \lesssim \delta.$$

This together with (3.24), $k_\delta \le \bar{k}_\delta$ and $\|r_{\alpha_k}(\mathscr{A})e_0\| \lesssim \alpha_k^\nu\|\omega\|$ then gives

$$\|e_{k_\delta}^\delta\| \lesssim \alpha_{k_\delta}^\nu\|\omega\| + \frac{\delta}{\sqrt{\alpha_{k_\delta}}} + \frac{\delta}{\sqrt{\alpha_{\bar{k}_\delta}}} \lesssim \alpha_{k_\delta}^\nu\|\omega\| + \frac{\delta}{\sqrt{\alpha_{\bar{k}_\delta}}}. \tag{3.35}$$

Using the definition of $\bar{k}_\delta$ and (1.4), we therefore complete the proof. □

4 Proof of Theorem 2

In this section we will give the proof of Theorem 2. The essential idea is similar as in the proof of Theorem 1. Thus we need to establish similar results as those used in Section 3. However, since we do not have source representation $e_0 = F'(x^\dagger)^*u$ any longer and since F satisfies different conditions, we must modify the arguments carefully. We will indicate the essential steps without spelling out all the necessary smallness conditions on $(K_0+K_1+K_2)\|e_0\|$. We first introduce the integer n_δ by

$$\alpha_{n_\delta} \le \left(\frac{\delta}{\gamma_1\|e_0\|}\right)^2 < \alpha_k, \qquad 0 \le k < n_\delta. \tag{4.1}$$

Recall that γ_1 is a constant satisfying $\gamma_1 > c_3 r^{1/2}/(\tau-1)$.

Proof of Theorem 2. In order to complete the proof of Theorem 2, we need to establish various estimates. We will divide the arguments into several steps.

Step 1. We will show that for all $0 \le k \le n_\delta$

$$x_k^\delta \in B_\rho(x^\dagger), \quad \|e_k^\delta\| \lesssim \|e_0\|, \tag{4.2}$$

$$\|F'(x^\dagger)e_k^\delta\| \lesssim \alpha_k^{1/2}\|e_0\| \tag{4.3}$$

and that $k_\delta \le n_\delta$ for the integer k_δ defined by the discrepancy principle (1.7) with $\tau > 1$.

To see this, we note that, for any $0 \le k < n_\delta$ with $x_k^\delta \in B_\rho(x^\dagger)$, (3.5) and Assumption 5 imply

$$\begin{aligned} e_{k+1}^\delta = {} & r_{\alpha_k}(\mathscr{A}_k^\delta)e_0 - \int_0^1 g_{\alpha_k}(\mathscr{A}_k^\delta)\mathscr{A}_k^\delta\left(R(x_k^\delta - te_k^\delta, x_k^\delta) - I\right)e_k^\delta dt \\ & + g_{\alpha_k}(\mathscr{A}_k^\delta)F'(x_k^\delta)^*(y^\delta - y). \end{aligned}$$

Therefore, with the help of Assumption 1(a) and (2.1), we have

$$\|e_{k+1}^\delta\| \le \|e_0\| + \frac{1}{2}K_0\|e_k^\delta\|^2 + c_4\delta\alpha_k^{-1/2} \le (1+c_4\gamma_1)\|e_0\| + \frac{1}{2}K_0\|e_k^\delta\|^2.$$

Thus, if $2(1+c_4\gamma_1)K_0\|e_0\| \le 1$, then, by using $\rho > 2(1+c_4\gamma_1)\|e_0\|$ and an induction argument, we can conclude $\|e_k^\delta\| \le 2(1+c_4\gamma_1)\|e_0\| < \rho$ for all $0 \le k \le n_\delta$. This establishes (4.2).

Next we show (4.3). It follows from (3.5), Assumption 1(a), (1.2), (2.19) and (4.1) that for $0 \le k < n_\delta$

$$\begin{aligned}\|F'(x_k^\delta)e_{k+1}^\delta\| &\lesssim \alpha_k^{1/2}\|e_0\| + \delta + \|F(x_k^\delta) - y - F'(x_k^\delta)e_k^\delta\| \\ &\lesssim \alpha_k^{1/2}\|e_0\| + (K_1+K_2)\|e_k^\delta\|\|F'(x^\dagger)e_k^\delta\|.\end{aligned}$$

By Assumption 6 we have

$$\|[F'(x^\dagger) - F'(x_k^\delta)]e_{k+1}^\delta\| \le K_1\|e_k^\delta\|\|F'(x^\dagger)e_{k+1}^\delta\| + K_2\|e_{k+1}^\delta\|\|F'(x^\dagger)e_k^\delta\|.$$

The above two inequalities and (4.2) then imply

$$\|F'(x^\dagger)e_{k+1}^\delta\| \lesssim \alpha_k^{1/2}\|e_0\| + K_1\|e_0\|\|F'(x^\dagger)e_{k+1}^\delta\| + (K_1+K_2)\|e_0\|\|F'(x^\dagger)e_k^\delta\|.$$

Thus, if $(K_1+K_2)\|e_0\|$ is sufficiently small, we can conclude (4.3) by an induction argument. As direct consequences of (4.2), (4.3) and Assumption 6 we have

$$\|F'(x_k^\delta)e_k^\delta\| \lesssim \alpha_k^{1/2}\|e_0\|, \quad 0 \le k \le n_\delta \tag{4.4}$$

and

$$\|F'(x_{k+1}^\delta)(x_{k+1}^\delta - x_k^\delta)\| \lesssim \alpha_k^{1/2}\|e_0\|, \quad 0 \le k < n_\delta. \tag{4.5}$$

In order to show $k_\delta \le n_\delta$, we note that (3.5) gives

$$\begin{aligned}F'(x^\dagger)e_{k+1}^\delta - y^\delta + y &= F'(x_k^\delta)r_{\alpha_k}(\mathscr{A}_k^\delta)e_0 + \left(F'(x^\dagger) - F'(x_k^\delta)\right) r_{\alpha_k}(\mathscr{A}_k^\delta)e_0 \\ &\quad - \left(F'(x^\dagger) - F'(x_k^\delta)\right) g_{\alpha_k}(\mathscr{A}_k^\delta)F'(x_k^\delta)^* \left(F(x_k^\delta) - y^\delta - F'(x_k^\delta)e_k^\delta\right) \\ &\quad - g_{\alpha_k}(\mathscr{B}_k^\delta)\mathscr{B}_k^\delta \left(F(x_k^\delta) - y - F'(x_k^\delta)e_k^\delta\right) - r_{\alpha_k}(\mathscr{B}_k^\delta)(y^\delta - y).\end{aligned}$$

Thus, by using (1.2), Assumption 1(a), (2.1), Assumption 6, (2.18), (4.2), (4.4) and (1.4) we have for $0 \le k < n_\delta$

$$\begin{aligned}\|F'(x^\dagger)e_{k+1}^\delta - y^\delta + y\| &\le \delta + c_3\alpha_k^{1/2}\|e_0\| + c_3K_1\|e_0\|\|e_k^\delta\|\alpha_k^{1/2} + K_2\|e_0\|\|F'(x_k^\delta)e_k^\delta\| \\ &\quad + K_1\|e_k^\delta\| \left(\delta + \frac{1}{2}(K_1+K_2)\|e_k^\delta\|\|F'(x_k^\delta)e_k^\delta\|\right) \\ &\quad + c_4K_2\alpha_k^{-1/2}\|F'(x_k^\delta)e_k^\delta\| \left(\delta + \frac{1}{2}(K_1+K_2)\|e_k^\delta\|\|F'(x_k^\delta)e_k^\delta\|\right) \\ &\quad + \frac{1}{2}(K_1+K_2)\|e_k^\delta\|\|F'(x_k^\delta)e_k^\delta\| \\ &\le \delta + (c_3 + C(K_1+K_2)\|e_0\|)\,\alpha_k^{1/2}\|e_0\| \\ &\le \delta + r^{1/2}(c_3 + C(K_1+K_2)\|e_0\|)\,\alpha_{k+1}^{1/2}\|e_0\|.\end{aligned}$$

Recall that $\gamma_1 > c_3r^{1/2}/(\tau-1)$. Thus, with the help of (4.2), (4.3) and the definition of n_δ, one can see that, if $(K_1+K_2)\|e_0\|$ is sufficiently small, then

$$\begin{aligned}\|F(x_{n_\delta}^\delta) - y^\delta\| &\le \|F(x_{n_\delta}^\delta) - y - F'(x^\dagger)e_{n_\delta}^\delta\| + \|F'(x^\dagger)e_{n_\delta}^\delta - y^\delta + y\| \\ &\le \delta + r^{1/2}(c_3 + C(K_1+K_2)\|e_0\|)\,\alpha_{n_\delta}^{1/2}\|e_0\| \\ &\quad + \frac{1}{2}(K_1+K_2)\|e_{n_\delta}^\delta\|\|F'(x^\dagger)e_{n_\delta}^\delta\| \\ &\le \delta + r^{1/2}(c_3 + C(K_1+K_2)\|e_0\|)\,\alpha_{n_\delta}^{1/2}\|e_0\| \\ &\le \delta + r^{1/2}(c_3 + C(K_1+K_2)\|e_0\|)\,\gamma_1^{-1}\delta \\ &\le \tau\delta.\end{aligned}$$

This implies $k_\delta \le n_\delta$.

Step 2. We will show, for the noise-free iterated solutions $\{x_k\}$, that for all $k\ge 0$

$$\|r_{\alpha_k}(\mathscr{A})e_0\| \lesssim \|e_k\| \lesssim \|r_{\alpha_k}(\mathscr{A})e_0\|, \tag{4.6}$$

$$\|e_k\| \lesssim \|e_{k+1}\| \lesssim \|e_k\| \tag{4.7}$$

and for all $0\le k\le l$

$$\|e_k\| \lesssim \|e_l\| + \frac{1}{\sqrt{\alpha_l}}\|F(x_k)-y\|. \tag{4.8}$$

In fact, from (3.13) and Assumption 5 it is easy to see that

$$\|e_{k+1} - r_{\alpha_k}(\mathscr{A}_k)e_0\| \le \frac{1}{2}K_0\|e_k\|^2. \tag{4.9}$$

If $2K_0\|e_0\|\le 1$, then by induction we can see that $\{x_k\}$ is well-defined and

$$\|e_k\| \le 2\|e_0\| \qquad \text{for all } k\ge 0. \tag{4.10}$$

This together with (4.9) and (2.20) gives

$$\|e_{k+1} - r_{\alpha_k}(\mathscr{A})e_0\| \lesssim \|[r_{\alpha_k}(\mathscr{A}_k) - r_{\alpha_k}(\mathscr{A})]e_0\| + K_0\|e_k\|^2 \lesssim K_0\|e_0\|\|e_k\|. \tag{4.11}$$

Thus, by Assumption 2 and the smallness of $K_0\|e_0\|$ we obtain (4.6) by induction. (4.7) is an immediate consequence of (4.11) and (4.6).

In order to show (4.8), we first consider the case $k>0$. Note that $x_k - x_l$ has a similar expression as in (3.30), so we may use (2.20), Assumption 5 and (4.10) to obtain

$$\begin{aligned}\|x_k - x_l\| &\lesssim \|r_{\alpha_{k-1}}(\mathscr{A})e_0 - r_{\alpha_{l-1}}(\mathscr{A})e_0\| + K_0\|e_0\|\left(\|e_{k-1}\| + \|e_{l-1}\|\right)\\ &\quad + K_0\|e_{k-1}\|^2 + K_0\|e_{l-1}\|^2\\ &\lesssim \|[r_{\alpha_{k-1}}(\mathscr{A}) - r_{\alpha_{l-1}}(\mathscr{A})]e_0\| + K_0\|e_0\|\left(\|e_{k-1}\| + \|e_{l-1}\|\right).\end{aligned} \tag{4.12}$$

By Lemma 1 with $x=e_0$, $\bar{x}=e_k$, $\alpha=\alpha_{l-1}$, $\beta=\alpha_{k-1}$ and $A=F'(x^\dagger)$, we have

$$\|[r_{\alpha_{k-1}}(\mathscr{A}) - r_{\alpha_{l-1}}(\mathscr{A})]e_0\| \lesssim \|r_{\alpha_{k-1}}(\mathscr{A})e_0 - e_k\| + \frac{1}{\sqrt{\alpha_{l-1}}}\|F'(x^\dagger)e_k\|.$$

With the help of (2.18), (4.10), and the smallness of $(K_1+K_2)\|e_0\|$, we have

$$\|F'(x^\dagger)e_k\| \le \|F(x_k)-y\| + \frac{1}{2}\|F'(x^\dagger)e_k\|. \tag{4.13}$$

Therefore $\|F'(x^\dagger)e_k\| \le 2\|F(x_k)-y\|$. This together with (4.11) and (4.7) then implies

$$\|[r_{\alpha_{k-1}}(\mathscr{A}) - r_{\alpha_{l-1}}(\mathscr{A})]e_0\| \lesssim K_0\|e_0\|\|e_k\| + \frac{1}{\sqrt{\alpha_l}}\|F(x_k)-y\|.$$

Combining this with (4.12) gives

$$\|x_k - x_l\| \lesssim K_0\|e_0\|\|e_k\| + \|e_l\| + \frac{1}{\sqrt{\alpha_l}}\|F(x_k)-y\|$$

which implies (4.8) if $K_0\|e_0\|$ is sufficiently small.

For the case $k=0$, we can assume $l\ge 1$. Since (4.8) is valid for $k=1$, we may use (4.7) to conclude that (4.8) is also true for $k=0$.

Step 3. We will show for all $k\ge 0$ that

$$\|F'(x^\dagger)e_k\| \lesssim \|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + \alpha_k^{1/2}\|r_{\alpha_k}(\mathscr{A})e_0\|. \tag{4.14}$$

To this end, first we may use the similar manner in deriving (4.3) to conclude

$$\|F'(x^\dagger)e_k\| \lesssim \alpha_k^{1/2}\|e_0\|. \tag{4.15}$$

Note that Assumption 6 and (4.10) imply

$$\begin{aligned}\|[F'(x^\dagger)-F'(x_k)]e_k\| &\le (K_1+K_2)\|e_k\|\|F'(x^\dagger)e_k\|\\ &\lesssim (K_1+K_2)\|e_0\|\|F'(x^\dagger)e_k\|.\end{aligned}$$

Therefore

$$\|F'(x_k)e_k\| \lesssim \|F'(x^\dagger)e_k\|. \tag{4.16}$$

In particular this implies

$$\|F'(x_k)e_k\| \lesssim \alpha_k^{1/2}\|e_0\|. \tag{4.17}$$

By using (3.33), (2.21), Assumption 6, (2.18) and Assumption 1(a) we obtain

$$\begin{aligned}\|F'(x^\dagger)e_{k+1}\| \lesssim\, & \|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + (K_0+K_1)\|e_0\|\|e_k\|\alpha_k^{1/2}\\ & + K_2\|e_0\|\left(\|F'(x^\dagger)e_k\|+\|F'(x_k)e_k\|\right)\\ & + (K_1+K_2)\|e_k\|\|F'(x_k)e_k\| + K_1(K_1+K_2)\|e_k\|^2\|F'(x_k)e_k\|\\ & + K_2(K_1+K_2)\|e_k\|\|F'(x_k)e_k\|^2\alpha_k^{-1/2}.\end{aligned}$$

Thus, with the help of (4.6), (4.15), (4.16), (4.17) and (4.10), we obtain

$$\|F'(x^\dagger)e_{k+1}\| \lesssim \|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + \alpha_k^{1/2}\|r_{\alpha_k}(\mathscr{A})e_0\| + K_2\|e_0\|\|F'(x^\dagger)e_k\|.$$

The estimates (4.14) thus follows by Assumption 2 and an induction argument if $K_2\|e_0\|$ is sufficiently small.

Step 4. Now we will establish some stability estimates. We will show for all $0 \le k \le n_\delta$ that

$$\|x_k^\delta - x_k\| \lesssim \frac{\delta}{\sqrt{\alpha_k}} \tag{4.18}$$

and

$$\|F(x_k^\delta)-F(x_k)-y^\delta+y\| \le (1+C(K_0+K_1+K_2)\|e_0\|)\,\delta. \tag{4.19}$$

In order to show (4.18), we use again the decomposition (3.18) for $x_{k+1}^\delta - x_{k+1}$. We still have $\|I_2\| \le c_4\delta/\sqrt{\alpha_k}$. By using (2.20) the term I_1 can be estimated as

$$\|I_1\| \lesssim K_0\|e_0\|\|x_k^\delta - x_k\|.$$

In order to estimate I_3, we note that Assumption 5 implies

$$\begin{aligned}I_3 =& \int_0^1 \left[g_{\alpha_k}(\mathscr{A}_k)\mathscr{A}_k - g_{\alpha_k}(\mathscr{A}_k^\delta)\mathscr{A}_k^\delta\right][R(x_k-te_k,x_k)-I]\,e_k dt\\ &+\int_0^1 g_{\alpha_k}(\mathscr{A}_k^\delta)F'(x_k^\delta)^*\left[F'(x_k^\delta)-F'(x_k)\right][R(x_k-te_k,x_k)-I]\,e_k dt\\ =& \int_0^1 \left[r_{\alpha_k}(\mathscr{A}_k^\delta)-r_{\alpha_k}(\mathscr{A}_k)\right][R(x_k-te_k,x_k)-I]\,e_k dt\\ &+\int_0^1 g_{\alpha_k}(\mathscr{A}_k^\delta)\mathscr{A}_k^\delta\left[I-R(x_k,x_k^\delta)\right][R(x_k-te_k,x_k)-I]\,e_k dt.\end{aligned}$$

Thus, by using (2.20) and (4.10), we obtain

$$\|I_3\| \lesssim K_0^2\|e_k\|^2\|x_k^\delta - x_k\| \lesssim K_0^2\|e_0\|^2\|x_k^\delta - x_k\|.$$

In order to estimate I_4, we again use Assumption 5 to write

$$\begin{aligned} I_4 &= g_{\alpha_k}(\mathscr{A}_k^\delta)F'(x_k^\delta)^*\left[F(x_k)-F(x_k^\delta)-F'(x_k^\delta)(x_k-x_k^\delta)\right] \\ &\quad + g_{\alpha_k}(\mathscr{A}_k^\delta)F'(x_k^\delta)^*\left[F'(x_k^\delta)-F'(x_k)\right]e_k \\ &= \int_0^1 g_{\alpha_k}(\mathscr{A}_k^\delta)\mathscr{A}_k^\delta\left[R(x_k^\delta+t(x_k-x_k^\delta),x_k^\delta)-I\right](x_k-x_k^\delta)dt \\ &\quad + g_{\alpha_k}(\mathscr{A}_k^\delta)\mathscr{A}_k^\delta\left[I-R(x_k,x_k^\delta)\right]e_k. \end{aligned}$$

Hence, we may use (4.2) and (4.10) to derive that

$$\|I_4\| \lesssim K_0\|x_k^\delta-x_k\|^2+K_0\|e_k\|\|x_k^\delta-x_k\| \lesssim K_0\|e_0\|\|x_k^\delta-x_k\|.$$

Combining the above estimates we obtain for $0\le k<n_\delta$

$$\|x_{k+1}^\delta-x_{k+1}\| \lesssim \frac{\delta}{\sqrt{\alpha_k}}+K_0\|e_0\|\|x_k^\delta-x_k\|.$$

Thus, if $K_0\|e_0\|$ is sufficiently small, we can obtain (4.18) immediately.

Next we show (4.19) by using (3.20). We still have (3.21). In order to estimate $\|F'(x_k^\delta)I_1\|$, $\|F'(x_k^\delta)I_3\|$ and $\|F'(x_k^\delta)I_4\|$, we note that Assumption 6, (4.10), (4.15) and (4.18) imply

$$\begin{aligned} &\|[F'(x_k)-F'(x^\dagger)](x_k^\delta-x_k)\| \\ &\qquad \le K_1\|e_k\|\|F'(x^\dagger)(x_k^\delta-x_k)\|+K_2\|F'(x^\dagger)e_k\|\|x_k^\delta-x_k\| \\ &\qquad \lesssim K_1\|e_0\|\|F'(x^\dagger)(x_k^\delta-x_k)\|+K_2\|e_0\|\delta, \end{aligned}$$

which in turn gives

$$\|F'(x_k)(x_k^\delta-x_k)\| \lesssim \|F'(x^\dagger)(x_k^\delta-x_k)\|+\delta. \tag{4.20}$$

Similarly, we have

$$\|F'(x_k^\delta)(x_k^\delta-x_k)\| \lesssim \|F'(x^\dagger)(x_k^\delta-x_k)\|+\delta. \tag{4.21}$$

Thus, by using (2.21), (4.18), (4.20) and (4.21) we have

$$\begin{aligned} \|F'(x_k^\delta)I_1\| &\lesssim (K_0+K_1)\|e_0\|\alpha_k^{1/2}\|x_k^\delta-x_k\| \\ &\qquad + K_2\|e_0\|\left(\|F'(x_k^\delta)(x_k^\delta-x_k)\|+\|F'(x_k)(x_k^\delta-x_k)\|\right) \\ &\lesssim (K_0+K_1+K_2)\|e_0\|\delta+K_2\|e_0\|\|F'(x^\dagger)(x_k^\delta-x_k)\|. \end{aligned}$$

Moreover, by employing (3.22), (2.20), Assumption 6, (2.18), (4.10), (4.17), (4.18) and (4.20), $\|F'(x_k^\delta)I_3\|$ can be estimated as

$$\begin{aligned} \|F'(x_k^\delta)I_3\| &\lesssim (K_0+K_1)\|x_k^\delta-x_k\|\|u_k\|+\alpha_k^{-1/2}K_2\|F'(x_k)(x_k^\delta-x_k)\|\|u_k\| \\ &\lesssim (K_0+K_1+K_2)(K_1+K_2)\|e_0\|^2\delta \\ &\quad + K_2(K_1+K_2)\|e_0\|^2\|F'(x^\dagger)(x_k^\delta-x_k)\|. \end{aligned}$$

while, by using Assumption 6, (2.18), (4.2), (4.10), (4.4), (4.18), (4.20) and (4.21), $\|F'(x_k^\delta)I_4\|$ can be estimated as

$$\begin{aligned} \|F'(x_k^\delta)I_4\| &\le \|F(x_k^\delta)-F(x_k)-F'(x_k)(x_k^\delta-x_k)\|+\|[F'(x_k^\delta)-F'(x_k)]e_k^\delta\| \\ &\lesssim (K_1+K_2)\|x_k^\delta-x_k\|\|F'(x_k)(x_k^\delta-x_k)\| \\ &\quad + K_1\|x_k^\delta-x_k\|\|F'(x_k^\delta)e_k^\delta\|+K_2\|F'(x_k^\delta)(x_k^\delta-x_k)\|\|e_k^\delta\| \\ &\lesssim (K_1+K_2)\|e_0\|\delta+(K_1+K_2)\|e_0\|\|F'(x^\dagger)(x_k^\delta-x_k)\|. \end{aligned}$$

Combining the above estimates we get

$$\begin{aligned}&\|F'(x_k^\delta)(x_{k+1}^\delta - x_{k+1}) - y^\delta + y\| \\ &\quad \le (1 + C(K_0 + K_1 + K_2)\|e_0\|)\delta + C(K_1 + K_2)\|e_0\|\|F'(x^\dagger)(x_k^\delta - x_k)\|. \end{aligned} \tag{4.22}$$

This in particular implies

$$\|F'(x_k^\delta)(x_{k+1}^\delta - x_{k+1})\| \lesssim \delta + (K_1 + K_2)\|e_0\|\|F'(x^\dagger)(x_k^\delta - x_k)\|.$$

On the other hand, similar to the derivation of (4.20), by Assumption 6, (4.2), (4.4) and (4.18) we have for $0 \le k < n_\delta$ that

$$\|F'(x^\dagger)(x_{k+1}^\delta - x_{k+1})\| \lesssim K_2\|e_0\|\delta + \|F'(x_k^\delta)(x_{k+1}^\delta - x_{k+1})\|.$$

Therefore

$$\|F'(x^\dagger)(x_{k+1}^\delta - x_{k+1})\| \lesssim \delta + (K_1 + K_2)\|e_0\|\|F'(x^\dagger)(x_k^\delta - x_k)\|.$$

Thus, if $(K_1 + K_2)\|e_0\|$ is small enough, then we can conclude

$$\|F'(x^\dagger)(x_k^\delta - x_k)\| \lesssim \delta, \qquad 0 \le k \le n_\delta. \tag{4.23}$$

Combining this with (4.22) gives for $0 \le k < n_\delta$

$$\|F'(x_k^\delta)(x_{k+1}^\delta - x_{k+1}) - y^\delta + y\| \le (1 + C(K_0 + K_1 + K_2)\|e_0\|)\,\delta. \tag{4.24}$$

Hence, by using (4.24), Assumption 6, (4.2), (4.5), (4.18), (4.21) and (4.23), we obtain for $0 \le k \le n_\delta$

$$\|F'(x_k^\delta)(x_k^\delta - x_k) - y^\delta + y\| \le (1 + C(K_0 + K_1 + K_2)\|e_0\|)\,\delta.$$

This together with (2.18), (4.2) and (4.10) implies (4.19).

Step 5. Now we are ready to complete the proof. By using the definition of k_δ, (4.19), (2.18) and (4.14) we have for $0 \le k < k_\delta$

$$\begin{aligned}\tau\delta &\le \|F(x_k^\delta) - y^\delta\| \le \|F(x_k^\delta) - F(x_k) - y^\delta + y\| + \|F(x_k) - y\| \\ &\le (1 + C(K_0 + K_1 + K_2)\|e_0\|)\delta + C\|F'(x^\dagger)e_k\| \\ &\le (1 + C(K_0 + K_1 + K_2)\|e_0\|)\,\delta + C\|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + C\alpha_k^{1/2}\|r_{\alpha_k}(\mathscr{A})e_0\|.\end{aligned}$$

Since $\tau > 1$, by assuming $(K_0 + K_1 + K_2)\|e_0\|$ is small enough, we can conclude for $0 \le k < k_\delta$ that

$$(\tau - 1)\delta \lesssim \|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + \alpha_k^{1/2}\|r_{\alpha_k}(\mathscr{A})e_0\|. \tag{4.25}$$

When $x_0 - x^\dagger$ satisfies (1.10) for some $\omega \in X$ and $0 < \nu \le \bar{\nu} - 1/2$, by using (4.25), (4.8), (4.6), (4.18), (4.19) and the definition of k_δ, we can employ the similar argument as in the last part of the proof of Theorem 1 to conclude (2.22).

When $x_0 - x^\dagger$ satisfies (1.11) for some $\omega \in X$ and $\mu > 0$, we have from Assumption 1(a) and (2.3) that

$$\|r_{\alpha_k}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + \alpha_k^{1/2}\|r_{\alpha_k}(\mathscr{A})e_0\| \le \left(c_0 b_{2\mu}^{1/2} + b_\mu\right)\alpha_k^{1/2}\left(-\ln(\alpha_k/(2\alpha_0))\right)^{-\mu}\|\omega\|.$$

This and (4.25) imply that there exists a constant $C_\mu > 0$ such that

$$(\tau - 1)\delta < C_\mu \alpha_k^{1/2}\left(-\ln(\alpha_k/(2\alpha_0))\right)^{-\mu}\|\omega\|, \quad 0 \le k < k_\delta.$$

If we introduce the integer $\hat{k}_\delta$ by

$$\alpha_{\hat{k}_\delta}^{1/2}\left(-\ln(\alpha_{\hat{k}_\delta}/(2\alpha_0))\right)^{-\mu} \le \frac{(\tau - 1)\delta}{C_\mu\|\omega\|} < \alpha_k^{1/2}\left(-\ln(\alpha_k/(2\alpha_0))\right)^{-\mu}, \quad 0 \le k < \hat{k}_\delta,$$

then $k_\delta \le \hat{k}_\delta$. Thus, by using (4.8), (4.18), (4.19), the definition of k_δ and the fact $\|e_k\| \lesssim \|r_{\alpha_k}(\mathscr{A})e_0\| \lesssim (-\ln(\alpha_k/(2\alpha_0)))^{-\mu}\|\omega\|$, we can use the similar manner in deriving (3.35) to get

$$\|e^\delta_{k_\delta}\| \lesssim \left(-\ln(\alpha_{\hat{k}_\delta}/(2\alpha_0))\right)^{-\mu}\|\omega\| + \frac{\delta}{\sqrt{\alpha_{\hat{k}_\delta}}} \lesssim \frac{\delta}{\sqrt{\alpha_{\hat{k}_\delta}}}. \tag{4.26}$$

By elementary argument we can show from (1.4) and the definition of $\hat{k}_\delta$ that there is a constant $c_\mu > 0$ such that

$$\alpha_{\hat{k}_\delta} \ge r^{-1}\alpha_{\hat{k}_\delta - 1} \ge c_\mu \left(\frac{\delta}{\|\omega\|}\right)^2 \left(1 + \left|\ln\frac{\delta}{\|\omega\|}\right|\right)^{2\mu}.$$

This together with (4.26) implies the estimate (2.23). □

5 Proof of Theorem 3

If $x_0 = x^\dagger$, then $k_\delta = 0$ and the result is trivial. Therefore, we will assume $x_0 \ne x^\dagger$. We define $\hat{k}_\delta$ to be the first integer such that

$$\|r_{\alpha_{\hat{k}_\delta}}(\mathscr{A})\mathscr{A}^{1/2}e_0\| + \alpha^{1/2}_{\hat{k}_\delta}\|r_{\alpha_{\hat{k}_\delta}}(\mathscr{A})e_0\| \le c\delta,$$

where the constant $c > 0$ is chosen so that we may apply Lemma 6 or (4.25) to conclude $k_\delta \le \hat{k}_\delta$. By (1.4), such $\hat{k}_\delta$ is clearly well-defined and is finite. Moreover, by a contradiction argument it is easy to show that

$$\hat{k}_\delta \to \infty \quad \text{as } \delta \to 0. \tag{5.1}$$

Now, under the conditions of Theorem 3 (i) we use Lemma 2, Lemma 4 and (3.24), while under the conditions of Theorem 3 (ii) we use (4.18), (4.19), (4.6) and (4.8), then from the definition of k_δ we have

$$\begin{aligned}\|e^\delta_{k_\delta}\| &\lesssim \|e_{k_\delta}\| + \frac{\delta}{\sqrt{\alpha_{k_\delta}}} \lesssim \|e_{k_\delta}\| + \frac{\delta}{\sqrt{\alpha_{\hat{k}_\delta}}}\\ &\lesssim \|e_{\hat{k}_\delta}\| + \frac{1}{\sqrt{\alpha_{\hat{k}_\delta}}}\left(\|F(x_{k_\delta}) - y\| + \delta\right)\\ &\lesssim \|r_{\alpha_{\hat{k}_\delta}}(\mathscr{A})e_0\| + \frac{\delta}{\sqrt{\alpha_{\hat{k}_\delta}}}\\ &\lesssim \frac{\delta}{\sqrt{\alpha_{\hat{k}_\delta}}}.\end{aligned} \tag{5.2}$$

We therefore need to derive the lower bound of $\alpha_{\hat{k}_\delta}$ under the conditions on e_0. We set for each $\alpha > 0$ and $0 \le \mu \le \bar{\nu}$

$$c_\mu(\alpha) := \left[\int_0^{1/2} \alpha^{-2\mu} r_\alpha(\lambda)^2 \lambda^{2\mu} d(E_\lambda\omega, \omega)\right]^{1/2},$$

where $\{E_\lambda\}$ denotes the spectral family generated by $\mathscr{A}$. It is easy to see for each $0 \le \mu < \bar{\nu}$ that $\alpha^{-2\mu}r_\alpha(\lambda)^2\lambda^{2\mu}$ is uniformly bounded for all $\alpha > 0$ and $\lambda \in [0, 1/2]$ and $\alpha^{-2\mu}r_\alpha(\lambda)^2\lambda^{2\mu} \to 0$ as $\alpha \to 0$ for all $\lambda \in (0, 1/2]$. Since $\omega \in \mathscr{N}(F'(x^\dagger))^\perp$, by the dominated convergence theorem we have for each $0 \le \mu < \bar{\nu}$

$$c_\mu(\alpha) \to 0 \quad \text{as } \alpha \to 0. \tag{5.3}$$

By the definition of $\hat{k}_\delta$, (1.4), Assumption 2, and the condition $e_0 = \mathscr{A}^\nu \omega$ we have

$$\begin{aligned}\delta &\lesssim \|r_{\alpha_{\hat{k}_\delta - 1}}(\mathscr{A})\mathscr{A}^{1/2} e_0\| + \alpha_{\hat{k}_\delta - 1} \|r_{\alpha_{\hat{k}_\delta - 1}}(\mathscr{A}) e_0\| \\ &\lesssim \|r_{\alpha_{\hat{k}_\delta}}(\mathscr{A})\mathscr{A}^{1/2} e_0\| + \alpha_{\hat{k}_\delta} \|r_{\alpha_{\hat{k}_\delta}}(\mathscr{A}) e_0\| \\ &\lesssim \alpha_{\hat{k}_\delta}^{\nu+1/2} \left(c_\nu(\alpha_{\hat{k}_\delta}) + c_{\nu+1/2}(\alpha_{\hat{k}_\delta})\right)\end{aligned}$$

This implies

$$\alpha_{\hat{k}_\delta} \geq \left(\frac{c\delta}{c_\nu(\alpha_{\hat{k}_\delta}) + c_{\nu+1/2}(\alpha_{\hat{k}_\delta})}\right)^{2/(1+2\nu)}. \tag{5.4}$$

Combining (5.2) and (5.4) gives

$$\|e^\delta_{k_\delta}\| \lesssim \left(c_\nu(\alpha_{\hat{k}_\delta}) + c_{\nu+1/2}(\alpha_{\hat{k}_\delta})\right)^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)}$$

Since $0 \leq \nu < \bar{\nu} - 1/2$, this together with (5.1) and (5.3) gives the desired conclusion.

6 Applications

In this section we will consider some specific methods defined by (1.3) by presenting several examples of $\{g_\alpha\}$. We will verify that those assumptions in Section 2 are satisfied for these examples.

6.1 **Example 1**

We first consider the function g_α given by

$$g_\alpha(\lambda) = \frac{(\alpha+\lambda)^m - \alpha^m}{\lambda(\alpha+\lambda)^m}, \tag{6.1}$$

where $m \geq 1$ is a fixed integer. This function arises from the iterated Tikhonov regularization of order m for linear ill-posed problems. Note that when $m = 1$, the corresponding method defined by (1.3) is exactly the iteratively regularized Gauss-Newton method (1.8). It is clear that the residual function corresponding to (6.1) is

$$r_\alpha(\lambda) = \frac{\alpha^m}{(\alpha+\lambda)^m}.$$

By elementary calculations it is easy to see that Assumption 1(a) and (b) are satisfied with $c_0 = (m-1)^{m-1}/m^m$ and $c_1 = m$. Moreover (2.1) is satisfied with

$$c_3 = \frac{1}{\sqrt{2m-1}}\left(\frac{2m-1}{2m}\right)^m \quad \text{and} \quad c_4 = \left(1 - \left(\frac{m+1}{m+3}\right)^m\right)\sqrt{m}.$$

By using the elementary inequality

$$1 - (1-t)^n \leq \sqrt{nt}, \qquad 0 \leq t \leq 1 \tag{6.2}$$

for any integer $n \geq 0$, we have for $0 < \alpha \leq \beta$ and $\lambda \geq 0$ that

$$r_\beta(\lambda) - r_\alpha(\lambda) = r_\beta(\lambda)\left[1 - \left(1 - \frac{\lambda/\alpha - \lambda/\beta}{1+\lambda/\alpha}\right)^m\right] \leq m^{1/2}\sqrt{\frac{\lambda}{\alpha}} r_\beta(\lambda).$$

This verifies Assumption 1(c) with $c_2 = m^{1/2}$. It is well-known that the qualification for g_α is $\bar{\nu} = m$ and (2.2) is satisfied with $d_\nu = (\nu/m)^\nu((m-\nu)/m)^{m-\nu} \le 1$ for each $0 \le \nu \le m$. For the sequence $\{\alpha_k\}$ satisfying (1.4), Assumption 2 is satisfied with $c_5 = r^m$.

In order to verify Assumption 4, we note that

$$
\begin{aligned}
& r_\alpha(A^*A) - r_\alpha(B^*B) \\
&\quad = \alpha^m \sum_{i=1}^m (\alpha I + A^*A)^{-i}[A^*(B-A) + (B^* - A^*)B](\alpha I + B^*B)^{-m-1+i}.
\end{aligned} \tag{6.3}
$$

Thus, by using the estimates

$$
\|(\alpha I + A^*A)^{-i}(A^*A)^\mu\| \le \alpha^{-i+\mu} \quad \text{for } i \ge 1 \text{ and } 0 \le \mu \le 1,
$$

we can verify (2.11), (2.12) and (2.13) easily.

Note also that $g_\alpha(\lambda) = \alpha^{-1}\sum_{i=1}^m \alpha^i(\alpha+\lambda)^{-i}$. We have, by using (2.12),

$$
\begin{aligned}
\|[g_\alpha(A^*A) - g_\alpha(B^*B)]B^*\| &\le \alpha^{-1}\sum_{i=1}^m \|\alpha^i[(\alpha I + A^*A)^{-i} - (\alpha I + B^*B)^{-i}]B^*\| \\
&\lesssim \alpha^{-1}\|A - B\|,
\end{aligned}
$$

which verifies (2.14).

Finally we verify Assumption 7 by assuming that F satisfies Assumption 5 and Assumption 6. We will use the abbreviation $F'_x := F'(x)$ for $x \in B_\rho(x^\dagger)$. With the help of (6.3) with $A = F'_x$ and $B = F'_z$, we obtain from Assumption 5 that

$$
\begin{aligned}
& \|r_\alpha(F'^*_x F'_x) - r_\alpha(F'^*_z F'_z)\| \\
&\quad \le \alpha^m \sum_{i=1}^m \|(\alpha I + F'^*_x F'_x)^{-i} F'^*_x F'_x [R(z,x) - I](\alpha I + F'^*_z F'_z)^{-m-1+i}\| \\
&\quad + \alpha^m \sum_{i=1}^m \|(\alpha I + F'^*_x F'_x)^{-i}[I - R(x,z)]^* F'^*_z F'_z(\alpha I + F'^*_z F'_z)^{-m-1+i}\| \\
&\quad \le \alpha^m \sum_{i=1}^m \alpha^{-i+1}\|I - R(z,x)\|\alpha^{-m-1+i} + \alpha^m \sum_{i=1}^m \alpha^{-i}\|I - R(x,z)\|\alpha^{-m+i} \\
&\quad \lesssim \|I - R(z,x)\| + \|I - R(x,z)\| \\
&\quad \lesssim K_0\|x - z\|
\end{aligned}
$$

which verifies (2.20). In order to show (2.21), we note that, for any $a \in X$ and $b \in Y$ satisfying $\|a\| = \|b\| = 1$, (6.3) implies

$$
\begin{aligned}
& (F'_x[r_\alpha(F'^*_x F'_x) - r_\alpha(F'^*_z F'_z)]a, b) \\
&\quad \le \alpha^m \sum_{i=1}^m \alpha^{-i+1}\|(F'_z - F'_x)(\alpha I + F'^*_z F'_z)^{-m-1+i}a\|\|b\| \\
&\quad + \alpha^m \sum_{i=1}^m \alpha^{-m-1/2+i}\|(F'_z - F'_x)(\alpha I + F'^*_x F'_x)^{-i}F'^*_x b\|\|a\|.
\end{aligned}
$$

Thus, by using Assumption 6, we have

$$
\begin{aligned}
&(F_x'[r_\alpha(F_x'^*F_x') - r_\alpha(F_z'^*F_z')]a, b) \\
&\quad \le \alpha^m \sum_{i=1}^m \alpha^{-i+1} K_1 \|x-z\| \|F_z'(\alpha I + F_z'^* F_z')^{-m-1+i} a\| \\
&\quad + \alpha^m \sum_{i=1}^m \alpha^{-i+1} K_2 \|F_z'(x-z)\| \|(\alpha I + F_z'^* F_z')^{-m-1+i} a\| \\
&\quad + \alpha^m \sum_{i=1}^m \alpha^{-m-1/2+i} K_1 \|x-z\| \|F_x'(\alpha I + F_x'^* F_x')^{-i} F_x'^* b\| \\
&\quad + \alpha^m \sum_{i=1}^m \alpha^{-m-1/2+i} K_2 \|F_x'(x-z)\| \|(\alpha I + F_x'^* F_x')^{-i} F_x'^* b\| \\
&\quad \lesssim K_1 \alpha^{1/2} \|x-z\| + K_2 \left(\|F_x'(x-z)\| + \|F_z'(x-z)\|\right).
\end{aligned}
$$

This verifies (2.21).

The above analysis shows that Theorem 1, Theorem 2 and Theorem 3 are applicable for the method defined by (1.3) and (1.7) with g_α given by (6.1). Thus we obtain the following result.

Corollary 1 *Let F satisfy (2.8) and (2.9), let $\{\alpha_k\}$ be a sequence of numbers satisfying (1.4), and let $\{x_k^\delta\}$ be defined by (1.3) with g_α given by (6.1) for some fixed integer $m \ge 1$. Let k_δ be the first integer satisfying (1.7) with $\tau > 1$.*

(i) If F satisfies Assumption 3 and if $x_0 - x^\dagger$ satisfies (1.10) for some $\omega \in X$ and $1/2 \le \nu \le m - 1/2$, then

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\nu \|\omega\|^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)}$$

provided $L\|u\| \le \eta_0$, where $u \in \mathscr{N}(F'(x^\dagger)^*)^\perp \subset Y$ is the unique element such that $x_0 - x^\dagger = F'(x^\dagger)^* u$, $\eta_0 > 0$ is a constant depending only on r, τ and m, and $C_\nu > 0$ is a constant depending only on r, τ, m and ν.

(ii) Let F satisfy Assumption 5 and Assumption 6, and let $x_0 - x^\dagger \in N(F'(x^\dagger))^\perp$. Then there exists a constant $\eta_1 > 0$ depending only on r, τ and m such that if $(K_0 + K_1 + K_2)\|x_0 - x^\dagger\| \le \eta_1$ then

$$\lim_{\delta \to 0} x_{k_\delta}^\delta = x^\dagger,$$

moreover, when $x_0 - x^\dagger$ satisfies (1.10) for some $\omega \in X$ and $0 < \nu \le m - 1/2$, then

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\nu \|\omega\|^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)}$$

for some constant $C_\nu > 0$ depending only on r, τ, m and ν; while when $x_0 - x^\dagger$ satisfies (1.11) for some $\omega \in X$ and $\mu > 0$, then

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\mu \|\omega\| \left(1 + \left|\ln \frac{\delta}{\|\omega\|}\right|\right)^{-\mu}$$

for some constant C_μ depending only on r, τ, m and μ.

Corollary 1 with $m = 1$ reproduces those convergence results in [3,8] for the iteratively regularized Gauss-Newton method (1.8) together with the discrepancy principle (1.7) under somewhat different conditions on F. Note that those results in [3,8] require τ be sufficiently large, while our result is valid for any $\tau > 1$. This less restrictive requirement on τ is important in numerical computations since the absolute error could increase with respect to τ. Moreover, when $x_0 - x^\dagger$ satisfies (1.10) with $\nu = 1/2$, Corollary 1 with $m = 1$ improves the corresponding result in [3], since we only need the Lipschitz condition on F' here.

Corollary 1 shows that the method defined by (1.3) and (1.7) with g_α given by (6.1) is order optimal for $0 < \nu \le m - 1/2$. However, we can not expect better rate of convergence than $O(\delta^{(2m-1)/(2m)})$ even if $x_0 - x^\dagger$ satisfies (1.10) with $m - 1/2 < \nu \le m$. An a posteriori stopping rule without such saturation has been studied in [9,10] for the iteratively regularized Gauss-Newton method (1.8).

6.2 **Example 2**

We consider the function g_α given by

$$g_\alpha(\lambda) = \sum_{i=0}^{[1/\alpha]} (1-\lambda)^i \tag{6.4}$$

which arises from the Landweber iteration applying to linear ill-posed problems. With such choice of g_α, the method (1.3) becomes

$$x_{k+1}^\delta = x_0 - \sum_{i=0}^{[1/\alpha_k]} \left(I - F'(x_k^\delta)^* F'(x_k^\delta)\right)^i F'(x_k^\delta)^* \left(F(x_k^\delta) - y^\delta - F'(x_k^\delta)(x_k^\delta - x_0)\right)$$

which is equivalent to the form

$$\begin{aligned}
&x_{k,0}^\delta = x_0,\\
&x_{k,i+1}^\delta = x_{k,i}^\delta - F'(x_k^\delta)^* \left(F(x_k^\delta) - y^\delta + F'(x_k^\delta)(x_{k,i}^\delta - x_k^\delta)\right), \qquad 0 \le i \le [1/\alpha_k],\\
&x_{k+1}^\delta = x_{k,[1/\alpha_k]+1}^\delta.
\end{aligned}$$

This method has been considered in [12] and is called the Newton-Landweber iteration.

Note that the corresponding residual function is

$$r_\alpha(\lambda) = (1-\lambda)^{[1/\alpha]+1}. \tag{6.5}$$

It is easy to see that Assumption 1(a), (b) and (2.1) hold with

$$c_0 = \frac{1}{2}, \quad c_1 = 2, \quad c_3 = \frac{\sqrt{2}}{3} \quad \text{and} \quad c_4 = \sqrt{2}.$$

Moreover, by (6.2) we have for any $0 < \alpha \le \beta$ that

$$r_\beta(\lambda) - r_\alpha(\lambda) = r_\beta(\lambda)\left(1 - (1-\lambda)^{[1/\alpha]-[1/\beta]}\right) \le \sqrt{\frac{\lambda}{\alpha}} r_\beta(\lambda).$$

This verifies Assumption 1(c) with $c_2 = 1$. It is well-known that the qualification of linear Landweber iteration is $\bar{\nu} = \infty$ and (2.2) is satisfied with $d_\nu = \nu^\nu$ for each $0 \le \nu < \infty$.

In order to verify Assumption 2, we restrict the sequence $\{\alpha_k\}$ to be of the form $\alpha_k := 1/n_k$, where $\{n_k\}$ is a sequence of positive integers such that

$$0 \le n_{k+1} - n_k \le q \quad \text{and} \quad \lim_{k\to\infty} n_k = \infty \tag{6.6}$$

for some $q \ge 1$. Then for $\lambda \in [0, 1/2]$ we have

$$r_{\alpha_k}(\lambda) = (1-\lambda)^{n_k - n_{k+1}} r_{\alpha_{k+1}}(\lambda) \le 2^q r_{\alpha_{k+1}}(\lambda).$$

Thus Assumption 2 is also true.

In order to verify Assumption 4, we will use some techniques from [7,12] and the following well-known estimates

$$\|(I - A^*A)^j (A^*A)^\nu\| \le \nu^\nu (j+\nu)^{-\nu}, \quad j \ge 0,\ \nu \ge 0 \tag{6.7}$$

for any bounded linear operator A satisfying $\|A\| \le 1$.

For any $\alpha > 0$, we set $k := [1/\alpha]$. Let A and B be any two bounded linear operators satisfying $\|A\|, \|B\| \le 1$. Then it follows from (6.5) that

$$r_\alpha(A^*A) - r_\alpha(B^*B) = \sum_{j=0}^{k} (I - A^*A)^j \left[A^*(B-A) + (B^* - A^*)B\right](I - B^*B)^{k-j}. \tag{6.8}$$

By using (6.7) we have

$$\|r_\alpha(A^*A) - r_\alpha(B^*B)\| \lesssim \sum_{j=0}^{k} \left((j+1)^{-1/2} + (k+1-j)^{-1/2}\right) \|A-B\|$$
$$\lesssim \sqrt{k}\|A-B\| \lesssim \frac{1}{\sqrt{\alpha}}\|A-B\|.$$

This verifies (2.11).

From (6.8) we also have $A\left[r_\alpha(A^*A) - r_\alpha(B^*B)\right]B^* = J_1 + J_2$, where

$$J_1 := \sum_{j=0}^{k} (I - AA^*)^j AA^*(B-A)(I - B^*B)^{k-j}B^*,$$
$$J_2 := \sum_{j=0}^{k} A(I - A^*A)^j (B^* - A^*)(I - BB^*)^{k-j}BB^*.$$

In order to verify (2.13), it suffices to show $\|J_1\| \lesssim (k+1)^{-1/2}\|A-B\|$ since the estimate on J_2 is exactly the same. We write $J_1 = J_1^{(1)} + J_2^{(2)}$, where

$$J_1^{(1)} := \sum_{j=0}^{[k/2]} (I - AA^*)^j AA^*(B-A)(I - B^*B)^{k-j}B^*,$$
$$J_1^{(2)} := \sum_{j=[k/2]+1}^{k} (I - AA^*)^j AA^*(B-A)(I - B^*B)^{k-j}B^*.$$

With the help of (6.7), we can estimate $J_1^{(2)}$ as

$$\|J_1^{(2)}\| \lesssim \sum_{j=[k/2]+1}^{k} (j+1)^{-1}(k+j-1)^{-1/2}\|A-B\|$$
$$\lesssim (k+1)^{-1}\sum_{j=0}^{k}(k+1-j)^{-1/2}\|A-B\| \lesssim (k+1)^{-1/2}\|A-B\|.$$

In order to estimate $J_1^{(1)}$, we use $AA^* = I - (I - AA^*)$ to rewrite it as

$$\begin{aligned} J_1^{(1)} = &\sum_{j=0}^{[k/2]} (I - AA^*)^j (B-A)(I - B^*B)^{k-j}B^* \\ &- \sum_{j=1}^{[k/2]+1} (I - AA^*)^j (B-A)(I - B^*B)^{k+1-j}B^* \\ =&(B-A)(I - B^*B)^k B^* - (I - AA^*)^{[k/2]+1}(B-A)(I - B^*B)^{k-[k/2]}B^* \\ &+ \sum_{j=1}^{[k/2]} (I - AA^*)^j (B-A)(I - B^*B)^{k-j}(B^*B)B^*. \end{aligned}$$

Thus, in view of (6.7), we obtain

$$\begin{aligned} \|J_1^{(1)}\| \lesssim &(k+1)^{-1/2}\|A-B\| + (k-[k/2]+1)^{-1/2}\|A-B\| \\ &+ \sum_{j=1}^{[k/2]} (k-j+1)^{-3/2}\|A-B\| \\ \lesssim &(k+1)^{-1/2}\|A-B\|. \end{aligned}$$

We thus verify (2.13). The verification of (2.12) can be done similarly.

Applying the estimate (2.12), we obtain

$$\begin{aligned}\|\,[g_\alpha(A^*A)-g_\alpha(B^*B)]\,B^*\| &\le \sum_{j=1}^{k}\|\left[(I-A^*A)^j-(I-B^*B)^j\right]B^*\|\\ &\lesssim k\|A-B\| \lesssim \frac{1}{\alpha}\|A-B\|,\end{aligned}$$

which verifies (2.14).

Finally we verify Assumption 7 by assuming that F satisfies Assumption 5 and Assumption 6. From (6.8) and Assumption 5 it follows that

$$\begin{aligned}r_\alpha(F_x'^*F_x')-r_\alpha(F_z'^*F_z') &= \sum_{j=0}^{k}(I-F_x'^*F_x')^jF_x'^*F_x'(R(z,x)-I)(I-F_z'^*F_z')^{k-j}\\ &\quad+\sum_{j=0}^{k}(I-F_x'^*F_x')^j(I-R(x,z))^*F_z'^*F_z'(I-F_z'^*F_z')^{k-j}.\end{aligned}$$

Thus we may use the argument in the verification of (2.13) to conclude

$$\|r_\alpha(F_x'^*F_x')-r_\alpha(F_z'^*F_z')\| \lesssim \|I-R(x,z)\|+\|I-R(z,x)\| \lesssim K_0\|x-z\|.$$

This verifies (2.20).

By using (6.8) and Assumption 5 we also have for any $w\in X$

$$F_x'[r_\alpha(F_x'^*F_x')-r_\alpha(F_z'^*F_z')]w = Q_1+Q_2+Q_3+Q_4,$$

where

$$\begin{aligned}Q_1 &= \sum_{j=0}^{[k/2]}(I-F_x'F_x'^*)^j(F_x'F_x'^*)(F_z'-F_x')(I-F_z'^*F_z')^{k-j}w,\\ Q_2 &= \sum_{j=[k/2]+1}^{k}(I-F_x'F_x'^*)^j(F_x'F_x'^*)(F_z'-F_x')(I-F_z'^*F_z')^{k-j}w,\\ Q_3 &= \sum_{j=0}^{[k/2]}(I-F_x'F_x'^*)^jF_x'(I-R(x,z))^*(F_z'^*F_z')(I-F_z'^*F_z')^{k-j}w,\\ Q_4 &= \sum_{j=[k/2]+1}^{k}(I-F_x'F_x'^*)^jF_x'(I-R(x,z))^*(F_z'^*F_z')(I-F_z'^*F_z')^{k-j}w.\end{aligned}$$

By employing (6.7) it is easy to see that

$$\|Q_3\| \lesssim \sum_{j=0}^{[k/2]}(j+1)^{-1/2}(k-j+1)^{-1}\|I-R(x,z)\|\|w\| \lesssim (k+1)^{-1/2}K_0\|x-z\|\|w\|.$$

With the help of (6.7) and Assumption 6, we have

$$\begin{aligned}\|Q_2\| &\lesssim \sum_{j=[k/2]+1}^{k}(j+1)^{-1}\|(F_z'-F_x')(I-F_z'^*F_z')^{k-j}w\|\\ &\lesssim K_1\|x-z\|\sum_{j=[k/2]+1}^{k}(j+1)^{-1}(k-j+1)^{-1/2}\|w\|\\ &\quad+K_2\|F_z'(x-z)\|\sum_{j=[k/2]+1}^{k}(j+1)^{-1}\|w\|\\ &\lesssim (k+1)^{-1/2}K_1\|x-z\|\|w\|+K_2\|F_z'(x-z)\|\|w\|.\end{aligned}$$

By using the argument in the verification of (2.13) and Assumption 6 we obtain

$$
\begin{aligned}
\|Q_1\| &\lesssim \|(F_z' - F_x')(I - F_z'^* F_z')^k w\| + \|(F_z' - F_x')(I - F_z'^* F_z')^{k-[k/2]} w\| \\
&\quad + \sum_{j=1}^{[k/2]} \|(F_z' - F_x')(I - F_z'^* F_z')^{k-j} (F_z'^* F_z') w\| \\
&\lesssim (k+1)^{-1/2} K_1 \|x - z\| \|w\| + K_2 \|F_z'(x - z)\| \|w\| \\
&\quad + \sum_{j=1}^{[k/2]} \Big(K_1 \|x - z\| (k - j + 1)^{-3/2} + K_2 \|F_z'(x - z)\| (k - j + 1)^{-1} \Big) \|w\| \\
&\lesssim (k+1)^{-1/2} K_1 \|x - z\| \|w\| + K_2 \|F_z'(x - z)\| \|w\|.
\end{aligned}
$$

Using Assumption 5 and the the similar argument in the verification of (2.13) we also have

$$
\|Q_4\| \lesssim (k+1)^{-1/2} \|I - R(x,z)\| \|w\| \lesssim (k+1)^{-1/2} K_0 \|x - z\| \|w\|.
$$

Combining the above estimates we thus obtain for any $w \in X$

$$
\begin{aligned}
&\|F_x' [r_\alpha(F_x'^* F_x') - r_\alpha(F_z'^* F_z')] w\| \\
&\qquad \lesssim (K_0 + K_1) \alpha^{1/2} \|x - z\| \|w\| + K_2 \|F_z'(x - z)\| \|w\|
\end{aligned}
$$

which implies (2.21).

Therefore, Theorem 1, Theorem 2 and Theorem 3 are applicable for the method defined by (1.3) and (1.7) with g_α given by (6.4).

The similar argument as above also applies to the situation where g_α is given by

$$
g_\alpha(\lambda) := \sum_{i=0}^{[1/\alpha]} (1 + \lambda)^{-i}
$$

which arise from the Lardy's method for solving linear ill-posed problems.

In summary, we obtain the following result.

Corollary 2 *Let F satisfy (2.8) and (2.9), and let $\{\alpha_k\}$ be a sequence given by $\alpha_k = 1/n_k$, where $\{n_k\}$ is a sequence of positive integers satisfying (6.6) for some $q \ge 1$. Let $\{x_k^\delta\}$ be defined by (1.3) with*

$$
g_\alpha(\lambda) = \sum_{i=0}^{[1/\alpha]} (1 - \lambda)^i \qquad \text{or} \qquad g_\alpha(\lambda) = \sum_{i=0}^{[1/\alpha]} (1 + \lambda)^{-i},
$$

and let k_δ be the first integer satisfying (1.7) with $\tau > 1$.

(i) If F satisfies Assumption 3, and if $x_0 - x^\dagger$ satisfies (1.10) for some $\omega \in X$ and $\nu \ge 1/2$, then

$$
\|x_{k_\delta}^\delta - x^\dagger\| \le C_\nu \|\omega\|^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)}
$$

provided $L\|u\| \le \eta_0$, where $u \in \mathcal{N}(F'(x^\dagger)^*)^\perp \subset Y$ is the unique element such that $x_0 - x^\dagger = F'(x^\dagger)^* u$, $\eta_0 > 0$ is a constant depending only on τ and q, and C_ν is a constant depending only on τ, q and ν.

(ii) Let F satisfy Assumption 5 and Assumption 6, and let $x_0 - x^\dagger \in N(F'(x^\dagger))^\perp$. Then there exists a constant $\eta_1 > 0$ depending only on τ and q such that if $(K_0 + K_1 + K_2)\|x_0 - x^\dagger\| \le \eta_1$ then

$$
\lim_{\delta \to 0} x_{k_\delta}^\delta = x^\dagger,
$$

moreover, when $x_0 - x^\dagger$ satisfies (1.10) for some $\omega \in X$ and $\nu > 0$, then

$$
\|x_{k_\delta}^\delta - x^\dagger\| \le C_\nu \|\omega\|^{1/(1+2\nu)} \delta^{2\nu/(1+2\nu)}
$$

for some constant $C_\nu > 0$ depending only on τ, q and ν; while when $x_0 - x^\dagger$ satisfies (1.11) for some $\omega \in X$ and $\mu > 0$, then

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\mu \|\omega\| \left(1 + \left|\ln \frac{\delta}{\|\omega\|}\right|\right)^{-\mu}$$

for some constant C_μ depending only on τ, q and μ.

6.3 **Example 3**

As the last example we consider the method (1.3) with g_α given by

$$g_\alpha(\lambda) = \frac{1}{\lambda}\left(1 - e^{-\lambda/\alpha}\right) \tag{6.9}$$

which arises from the asymptotic regularization for linear ill-posed problems. In this method, the iterated sequence $\{x_k^\delta\}$ is equivalently defined as $x_{k+1}^\delta := x^\delta(1/\alpha_k)$, where $x^\delta(t)$ is the solution of the initial value problem

$$\begin{aligned} &\frac{d}{dt}x^\delta(t) = F'(x_k^\delta)^*\left(y^\delta - F(x_k^\delta) + F'(x_k^\delta)(x_k^\delta - x^\delta(t))\right), \quad t > 0,\\ &x^\delta(0) = x_0. \end{aligned}$$

Note that the corresponding residual function is

$$r_\alpha(\lambda) = e^{-\lambda/\alpha}.$$

It is easy to see that Assumption 1(a), (b) and (2.1) hold with

$$c_0 = e^{-1}, \quad c_1 = 1, \quad c_3 = \frac{1}{\sqrt{2e}} \quad \text{and} \quad c_4 = \sqrt{\frac{2}{e}}.$$

By using the inequality $1 - e^{-t} \le \sqrt{t}$ for $t \ge 0$ we have for $0 < \alpha \le \beta$ that

$$r_\beta(\lambda) - r_\alpha(\lambda) = r_\beta(\lambda)\left(1 - e^{\lambda/\beta - \lambda/\alpha}\right) \le \sqrt{\frac{\lambda}{\alpha} - \frac{\lambda}{\beta}}\, r_\beta(\lambda) \le \sqrt{\frac{\lambda}{\alpha}}\, r_\beta(\lambda).$$

This verifies Assumption 1(c) with $c_2 = 1$. It is well-known that the qualification of the linear asymptotic regularization is $\bar{\nu} = \infty$ and (2.2) is satisfied with $d_\nu = (\nu/e)^\nu$ for each $0 \le \nu < \infty$.

In order to verify Assumption 2, we assume that $\{\alpha_k\}$ is a sequence of positive numbers satisfying

$$0 \le \frac{1}{\alpha_{k+1}} - \frac{1}{\alpha_k} \le \theta_0 \quad \text{and} \quad \lim_{k\to\infty} \alpha_k = 0 \tag{6.10}$$

for some $\theta_0 > 0$. Then for all $\lambda \in [0,1]$ we have

$$r_{\alpha_k}(\lambda) = e^{(1/\alpha_{k+1} - 1/\alpha_k)\lambda} r_{\alpha_{k+1}}(\lambda) \le e^{\theta_0} r_{\alpha_{k+1}}(\lambda).$$

Thus Assumption 2 is also true.

In order to verify Assumption 4 and Assumption 7, we set for every integer $n \ge 1$

$$r_{\alpha,n}(\lambda) := \left(1 + \frac{\lambda}{n\alpha}\right)^{-n}, \qquad g_{\alpha,n}(\lambda) := \frac{1}{\lambda}\left(1 - \left(1 + \frac{\lambda}{n\alpha}\right)^{-n}\right).$$

Note that, for each fixed $\alpha > 0$, $\{r_{\alpha,n}\}$ and $\{g_{\alpha,n}\}$ are uniformly bounded over $[0,1]$, and $r_{\alpha,n}(\lambda) \to r_\alpha(\lambda)$ and $g_{\alpha,n}(\lambda) \to g_\alpha(\lambda)$ as $n \to \infty$. By the dominated convergence theorem, we have for any bounded linear operator A with $\|A\| \le 1$ that

$$\begin{aligned}
&\lim_{n\to\infty} \|[r_\alpha(A^*A) - r_{\alpha,n}(A^*A)]x\|^2 \\
&\qquad = \lim_{n\to\infty} \int_0^{\|A\|^2} (r_\alpha(\lambda) - r_{\alpha,n}(\lambda))^2 \, d(E_\lambda x, x) = 0
\end{aligned}$$

and

$$\begin{aligned}
&\lim_{n\to\infty} \|[g_\alpha(A^*A) - g_{\alpha,n}(A^*A)]x\|^2 \\
&\qquad = \lim_{n\to\infty} \int_0^{\|A\|^2} (g_\alpha(\lambda) - g_{\alpha,n}(\lambda))^2 \, d(E_\lambda x, x) = 0
\end{aligned}$$

for any $x \in X$, where $\{E_\lambda\}$ denotes the spectral family generated by A^*A. Thus it suffices to verify Assumption 4 and Assumption 7 with g_α and r_α replaced by $g_{\alpha,n}$ and $r_{\alpha,n}$ with uniform constants c_6, c_7 and c_8 independent of n. Let A and B be any two bounded linear operators satisfying $\|A\|, \|B\| \le 1$. We need the following inequality which says for any integer $n \ge 1$ there holds

$$\|r_{\alpha,n}(A^*A)(A^*A)^\nu\| \le \nu^\nu \alpha^\nu, \qquad 0 \le \nu \le n. \tag{6.11}$$

By noting that

$$\begin{aligned}
&r_{\alpha,n}(A^*A) - r_{\alpha,n}(B^*B) \\
&\quad = \frac{1}{n\alpha} \sum_{i=1}^n r_{\alpha,i}(A^*A)\,[A^*(B-A) + (B^* - A^*)B]\, r_{\alpha,n+1-i}(B^*B),
\end{aligned} \tag{6.12}$$

we thus obtain

$$\begin{aligned}
&\|r_{\alpha,n}(A^*A) - r_{\alpha,n}(B^*B)\| \le \sqrt{\frac{2}{\alpha}} \|A - B\|, \\
&\|[r_{\alpha,n}(A^*A) - r_{\alpha,n}(B^*B)]B^*\| \le \frac{3}{2} \|A - B\|
\end{aligned} \tag{6.13}$$

and

$$\|A[r_{\alpha,n}(A^*A) - r_{\alpha,n}(B^*B)]B^*\| \le \sqrt{2\alpha} \|A - B\|.$$

Furthermore, by noting that $g_{\alpha,n}(\lambda) = \frac{1}{n\alpha} \sum_{i=1}^n r_{\alpha,i}(\lambda)$, we may use (6.13) to conclude

$$\begin{aligned}
\|[g_{\alpha,n}(A^*A) - g_{\alpha,n}(B^*B)]B^*\| &\le \frac{1}{n\alpha} \sum_{i=1}^n \|[r_{\alpha,i}(A^*A) - r_{\alpha,i}(B^*B)]B^*\| \\
&\le \frac{3}{2\alpha} \|A - B\|.
\end{aligned}$$

Assumption 4 is therefore verified.

It remains to verify Assumption 7 with g_α and r_α replaced by $g_{\alpha,n}$ and $r_{\alpha,n}$ with uniform constants c_7 and c_8 independent of n. By using (6.12), Assumption 5 and (6.11) we have

$$\begin{aligned}
&\|r_{\alpha,n}(F_x'^* F_x') - r_{\alpha,n}(F_z'^* F_z')\| \\
&\quad \le \frac{1}{n\alpha} \sum_{i=1}^n \|r_{\alpha,i}(F_x'^* F_x')(F_x'^* F_x')(R(z,x) - I) r_{\alpha,n+1-i}(F_z'^* F_z')\| \\
&\quad + \frac{1}{n\alpha} \sum_{i=1}^n \|r_{\alpha,i}(F_x'^* F_x')(I - R(x,z))^* (F_z'^* F_z') r_{\alpha,n+1-i}(F_z'^* F_z')\| \\
&\quad \le \|I - R(z,x)\| + \|I - R(x,z)\| \\
&\quad \le 2K_0 \|x - z\|.
\end{aligned}$$

This implies (2.20).

By using (6.12), Assumption 6 and (6.11) we also have for any $a \in X$ and $b \in Y$ satisfying $\|a\| = \|b\| = 1$ that

$$
\begin{aligned}
&(F_x'[r_{\alpha,n}(F_x'^*F_x') - r_{\alpha,n}(F_z'^*F_z')]a,b) \\
&\quad \le \frac{1}{n\alpha}\sum_{i=1}^{n} |(r_{\alpha,i}(F_x'F_x'^*)(F_x'F_x'^*)(F_z'-F_x')r_{\alpha,n+1-i}(F_z'^*F_z')a,b)| \\
&\quad + \frac{1}{n\alpha}\sum_{i=1}^{n} |(a, r_{\alpha,n+1-i}(F_z'^*F_z')F_z'^*(F_z'-F_x')F_x'^*r_{\alpha,i}(F_x'F_x'^*)b)| \\
&\quad \le \sqrt{2}K_1\alpha^{1/2}\|x-z\| + K_2\|F_z'(x-z)\| + \frac{1}{2}K_2\|F_x'(x-z)\|.
\end{aligned}
$$

This implies (2.21).

Therefore, we may apply Theorem 1, Theorem 2 and Theorem 3 to conclude the following result.

Corollary 3 *Let F satisfy (2.8) and (2.9), and let $\{\alpha_k\}$ be a sequence of positive numbers satisfying (6.10) for some $\theta_0 > 0$. Let $\{x_k^\delta\}$ be defined by (1.3) with g_α given by (6.9) and let k_δ be the first integer satisfying (1.7) with $\tau > 1$.*

(i) If F satisfies Assumption 3, and if $x_0 - x^\dagger$ satisfies (1.10) for some $\omega \in X$ and $\nu \ge 1/2$, then

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\nu \|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}$$

provided $L\|u\| \le \eta_0$, where $u \in \mathscr{N}(F'(x^\dagger)^*)^\perp \subset Y$ is the unique element such that $x_0 - x^\dagger = F'(x^\dagger)^*u$, $\eta_0 > 0$ is a constant depending only on τ, θ_0 and α_0, and C_ν is a constant depending only on τ, θ_0, α_0 and ν.

(ii) Let F satisfy Assumption 5 and Assumption 6, and let $x_0 - x^\dagger \in N(F'(x^\dagger))^\perp$. Then there exists a constant $\eta_1 > 0$ depending only on τ, θ_0 and α_0 such that if $(K_0+K_1+K_2)\|x_0 - x^\dagger\| \le \eta_1$ then

$$\lim_{\delta\to 0} x_{k_\delta}^\delta = x^\dagger;$$

moreover, when $x_0 - x^\dagger$ satisfies (1.10) for some $\omega \in X$ and $\nu > 0$, then

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\nu \|\omega\|^{1/(1+2\nu)}\delta^{2\nu/(1+2\nu)}$$

for some constant $C_\nu > 0$ depending only on τ, θ_0, α_0 and ν; while when $x_0 - x^\dagger$ satisfies (1.11) for some $\omega \in X$ and $\mu > 0$, then

$$\|x_{k_\delta}^\delta - x^\dagger\| \le C_\mu \|\omega\| \left(1 + \left|\ln \frac{\delta}{\|\omega\|}\right|\right)^{-\mu}$$

for some constant C_μ depending only on τ, θ_0, α_0 and μ.

Acknowledgements The authors wish to thank the referee for careful reading of the manuscript and useful comments.

References

1. A. B. Bakushinskii, *The problems of the convergence of the iteratively regularized Gauss-Newton method*, Comput. Math. Math. Phys., 32(1992), 1353–1359.
2. F. Bauer and T. Hohage, *A Lepskij-type stopping rule for regularized Newton methods*, Inverse Problems, 21(2005), 1975–1991.
3. B. Blaschke, A. Neubauer and O. Scherzer, *On convergence rates for the iteratively regularized Gauss-Newton method*, IMA J. Numer. Anal., 17(1997), 421–436.
4. P. Deuflhard, H. W. Engl and O. Scherzer, *A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions*, Inverse Problems, 14 (1998), no. 5, 1081–1106.

5. M. Hanke, *A regularizing Levenberg-Marquardt scheme with applications to inverse groundwater filtration problems*, Inverse Problems, 13(1997), 79–95.
6. M. Hanke, *Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse problems*, Numer. Funct. Anal. Optim., 18(1997), 971–993.
7. M. Hanke, A. Neubauer and O. Scherzer, *A convergence analysis of Landweber iteration of nonlinear ill-posed problems*, Numer. Math., 72(1995), 21–37.
8. T. Hohage, *Logarithmic convergence rates of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem*, Inverse Problems, 13 (1997), no. 5, 1279–1299.
9. Q. N. Jin, *On the iteratively regularized Gauss-Newton method for solving nonlinear ill-posed problems*, Math. Comp., 69 (2000), no. 232, 1603–1623.
10. Q. N. Jin, *A convergence analysis of the iteratively regularized Gauss-Newton method under the Lipschitz condition*, Inverse Problems, 24(2008), no. 4, to appear.
11. Q. N. Jin and Z. Y. Hou, *On an a posteriori parameter choice strategy for Tikhonov regularization of nonlinear ill-posed problems*, Numer. Math., 83(1999), no. 1, 139–159.
12. B. Kaltenbacher, *Some Newton-type methods for the regularization of nonlinear illposed problems*, Inverse Problems, 13(1997), 729–753.
13. B. Kaltenbacher, *A posteriori choice strategies for some Newton type methods for the regularization of nonlinear ill-posed problems*, Numer. Math., 79 (1998), 501-528.
14. B. Kaltenbacher, A. Neubauer and O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, Berlin, de Gruyter, 2008.
15. A. Rieder, *On the regularization of nonlinear ill-posed problems via inexact Newton iterations*, Inverse Problems, 15(1999), 309–327.
16. A. Rieder, *On convergence rates of inexact Newton regularizations*, Numer. Math., 88(2001), 347–365.
17. O. Scherzer, H. W. Engl and K. Kunisch, *Optimal a posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems*, SIAM J. Numer. Anal., 30(1993), 1796–1838.
18. U. Tautenhahn, *On a general regularization scheme for nonlinear ill-posed problems*, Inverse Problems, 13 (1997), no. 5, 1427–1437.
19. U. Tautenhahn and Q. N. Jin, *Tikhonov regularization and a posteriori rules for solving nonlinear ill posed problems*, Inverse Problems, 19 (2003), no. 1, 1–21.
20. G. M. Vainikko and A. Y. Veretennikov, *Iteration Procedures in Ill-Posed Problems*, Moscow, Nauka, 1986 (In Russian).